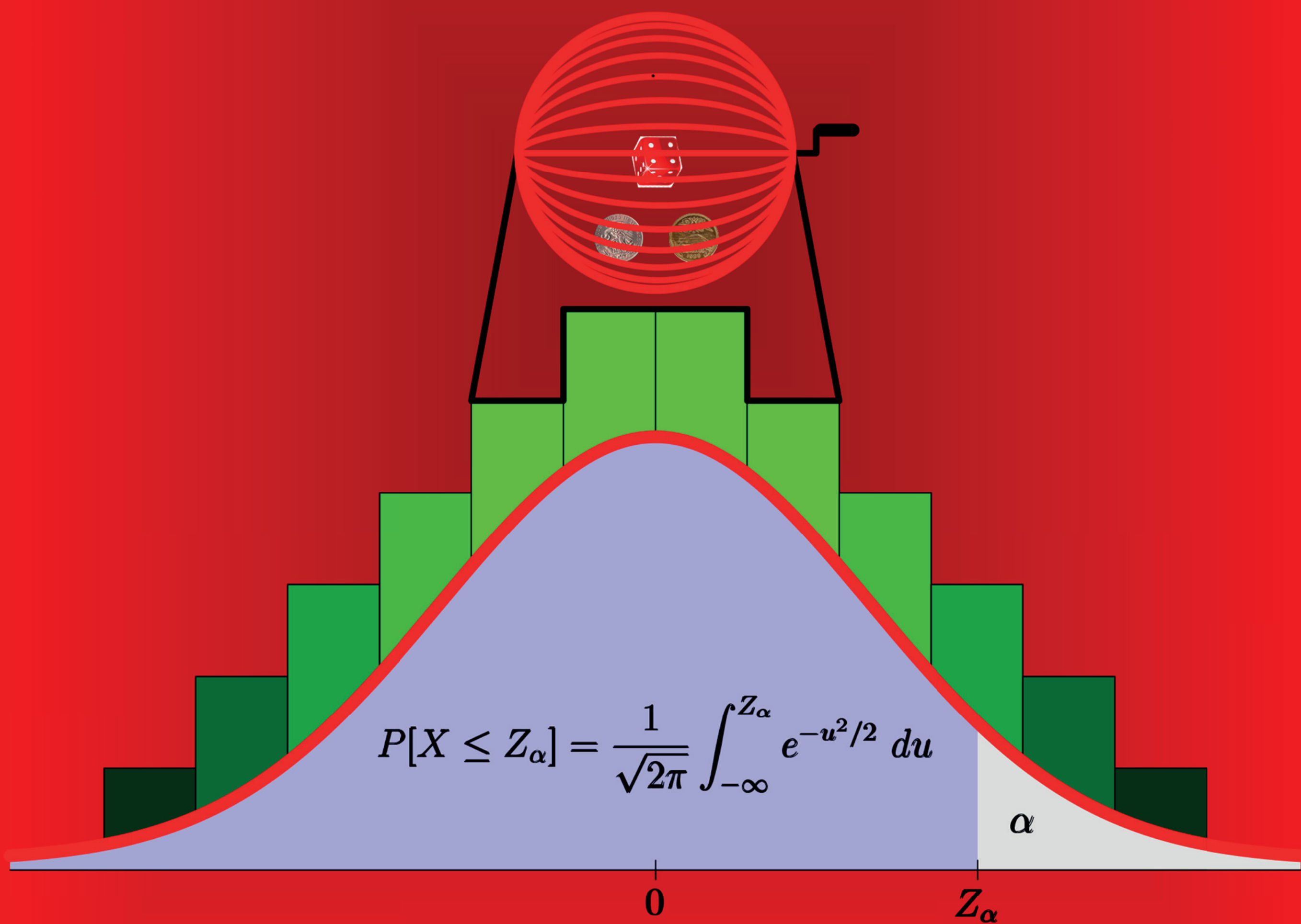
Estadística y probabilidad



Manuel Jesús Tec Canché

Estadística y probabilidad

Universidad Autónoma de la Ciudad de México

M en C. Juan Carlos Aguilar Franco
Rector

Dra. María Elizbeth Álvarez Sánchez
Coordinadora Académica

Museógrafo Fernando Fco. Félix y Valenzuela
Coordinador de Difusión Cultural y Extensión Universitaria

Equipo de la Biblioteca del Estudiante

Ángeles Godínez Guevara
Responsable

Ana Beatriz Alonso Osorio
Ana Lina Graciano
Daniel Valentin Cruz
Florina Piña Cancino
Heber Blass Bautista
Sergio Javier Cortés Becerril

Estadística y probabilidad

Manuel Jesús Tec Canché

UACM
Universidad Autónoma
de la Ciudad de México
NADA HUMANO ME ES AJENO

Biblioteca
BE
del
Estudiante

FICHA CATALOGRÁFICA E-S/N

Tec Canché, Manuel Jesús

Estadística y probabilidad / Manuel Jesús Tec Canché. -- Primera edición. -- Ciudad de México : Universidad Autónoma de la Ciudad de México, 2024.

vi, 123 páginas : gráficas, tablas ; 21 cm.

Bibliografía: página 123.

ISBN: 978-607-2615-48-9

1.Combinaciones (Matemáticas) 2. Análisis combinatorio - Problemas, ejercicios, etc. 3. Variables (Matemáticas) 4. Variables aleatorias - Problemas, ejercicios, etc. 5. Estadística - Problemas, ejercicios, etc. I. Título

LC QA164.5 Dewey 511.6

Estadística y probabilidad

primera edición, 2025

ISBN: 978-607-2615-48-9

https://www.uacm.edu.mx/Organizacion/CoordinacionAcademica/Biblioteca_Estudiante

Hecho e impreso en México

Índice general

Introducción V

1. Análisis combinatorio 1

1.1. Principio Fundamental de Conteo 1

1.2. Permutaciones y Combinaciones 1

1.2.1. Diagrama de árbol 2

1.3. Permutaciones y Combinaciones sin repetición 3

1.3.1. Permutaciones de Bloques ordenados 5

1.3.2. Combinaciones de varios tipos 6

1.4. Permutaciones y combinaciones con repetición 7

1.4.1. Permutación de palabras 7

1.4.2. Combinaciones con repetición 10

1.5. Ejercicios 12

2. Conceptos y Teoremas básicos de probabilidad 19

2.1. Elementos básicos 19

2.1.1. Experimento aleatorio y Espacio muestral 19

2.1.2. Espacio de Probabilidad 20

2.1.3. Espacio Muestral Finito 22

2.1.4. Independencia y Probabilidad Condicional 24

2.2. Teoremas Básicos 27

2.2.1. Teorema de la Multiplicación 27

2.2.2. Teorema de la Probabilidad Total 28

2.2.3. Teorema de Bayes 29

2.3. Ejercicios 31

3. Variables aleatorias 37

3.1. Función de distribución acumulada (fda) 37

3.1.1. Propiedades de la Función de distribución acumulada 38

3.2. Variable aleatoria Discreta 39

3.2.1. Función de distribución de probabilidad (fdp) 40

3.3. Variable aleatoria Continua 40

3.3.1. Función de densidad de probabilidad (fdp) 40

3.4. Variable aleatoria Mixta 41

3.5. Esperanza de una variable aleatoria 42

3.5.1. Propiedades del Valor esperado 42

3.6. Varianza de una variable aleatoria 43

3.6.1. Propiedades de la Varianza 43

3.7. Ejercicios 49

4. Variables aleatorias discretas de uso común 51

4.1. Distribución Bernoulli 51

4.2. Distribución Binomial 52

4.3. Distribución Hipergeométrica 56

4.4. Distribución Geométrica 59

4.5. Distribución Pascal o Binomial negativa 60

4.6. Distribución Poisson 61

4.7. Ejercicios 68

5. Variables aleatorias continuas de uso común 69

5.1. Distribución Uniforme 69

5.2. Distribución Normal 70

5.2.1. Aproximación Normal a la distribución Binomial 73

5.3. Distribución Gamma 74

5.3.1. Distribución Exponencial 75

5.3.2. Distribución Erlang 76

5.3.3. Distribución Ji-cuadrada 79

5.4. Ejercicios 81

6. Estadística descriptiva **83**

6.1. Clasificación de datos 83

6.2. Representación de datos 84

6.3. Medidas de tendencia central 84

6.4. Medidas de dispersión 86

6.5. Gráficas de Barra y de Pastel 86

6.6. Datos agrupados 88

6.7. Histograma, Polígono de frecuencias y Ojiva 93

6.8. Ejercicios 95

7. Estadística inferencial **97**

7.1. Estimación paramétrica 97

7.1.1. Estimación puntual 98

7.1.2. Características de un buen estimador 98

7.1.3. Estimación por intervalo de confianza 99

7.2. Pruebas de hipótesis paramétricas 102

7.3. Ejercicios 107

A. Tablas estadísticas **111**

A.1. Distribución normal estandar Z 111

A.1.1. Función: double pnorm(double x) 111

A.1.2. Función: double qnorm(double x) 112

A.1.3. Tabla de la distribución Normal estándar 112

A.2. Distribución t 113

A.2.1. Función: double pt(double x,int k) 113

A.2.2. Función: double qt(double x,int k) 113

A.2.3. Tabla de la distribución t 114

A.3. Distribución *ji-cuadrada* 114

A.3.1. Función: double pchisq(double x,int k); 114

A.3.2. Función: double qchisq(double y,int k) 116
A.3.3. Tabla de la distribución *ji-cuadrada* 116

Bibliography **121**

Introducción

Este libro se escribió con el propósito de apoyar a los estudiantes de Ingeniería y de Modelación Matemática en su curso de Estadística y Probabilidad de la Universidad Autónoma de la Ciudad de México.
La Estadística es una de las ramas de las matemáticas con una gran variedad de aplicaciones en diferentes áreas de las ciencias.
En este libro profundizaremos en la parte de probabilidad ya que la probabilidad es el lenguaje de la estadística.
Este material ha sido desarrollado por el autor como resultado de la enseñanza de la misma durante muchos años en las siguientes universidades: Universidad Autónoma Metropolitana Iztapalapa, Universidad Anáhuac Norte y Universidad Autónoma de la Ciudad de México.
El capítulo 1 aborda los conceptos básicos del análisis combinatorio: combinación y permutación.
En el capítulo 2 se dan los conceptos y resultados básicos de la probabilidad, y se empieza con los primeros ejemplos de experimentos aleatorios sobre espacios muestrales finitos cuyos elementos se trabajarán como combinaciones o permutaciones, conceptos trabajados en el capítulo 1.
En el capítulo 3 se estudia el concepto de variable aleatoria.
En capítulo 4 y capítulo 5 se estudian las variables aleatorias discretas y contínuas de uso común, las que son estudiadas en todo curso de probabildad y aparecen en una gran cantidad de aplicaciones reales importantes.
En el capítulo 6 correspondiente a estadística descriptiva se empieza a relacionar los conceptos de probabilidad a la estadística, por ejemplo el concepto de variable aleatoria con población, el de variable aleatoria con el de media aritmética, distribución de probabilidad con el de histograma, etc.
En el capítulo 7 empezamos con la estadística inferencial con los conceptos de estimación y la prueba de hipótesis paramétrica.
En el apéndice se construye con C++ las tablas de las siguientes tres distribuciones: $normal$, t, y $ji - cuadrada$.
Utilizamos el paquete estadístico R como herramienta de apoyo.

Finalmente quiero agradecer a la Universidad Autónoma de la Ciudad de México UACM su gran apoyo para poder llevar a cabo este proyecto UACM/CCT/CAS/037/18 realizado durante el año sabático 2019.

Ciudad de México 2024 Manuel Jesús Tec Canché.

Capítulo 1

Análisis combinatorio

En esta sección se estudian las **técnicas básicas de conteo** de ciertos conjuntos llamados combinaciones y permutaciones, que serán utilizadas en el siguiente capítulo para el cálculo de probabilidades sobre conjuntos finitos (ver [1], [2]).

1.1. Principio Fundamental de Conteo

Postulado 1.1.1 (Principio de la Multiplicación o Principio Fundamental de Conteo)**.** *La cantidad de formas posibles de realizar una tarea que consiste de r etapas separadas y donde la etapa i se realiza de n_i formas (para $i = 1, 2, \ldots r$), viene dada por el producto $n_1 \cdot n_2 \ldots n_r$.*
De ahora en adelante, cuando utilicemos el Principio de la Multiplicación y lo amerite, haremos mención a la siguiente tabla.

Etapa	1	2	$\cdots$	i	$\cdots$	r
Cantidad de formas en que se realiza la etapa	n_1	n_2	$\cdots$	n_i	$\cdots$	n_r

1.2. Permutaciones y Combinaciones

Definición 1.2.1. *Una permutación de r objetos $s_1, \ldots, s_r$, es un arreglo* ***ordenado*** *de dichos objetos, la cual se representa como $(s_1 s_2 \ldots s_r)$, en donde se indica que para $i = 1, \ldots, r$ el i-ésimo elemento de la permutación es el objeto s_i.*

Ejemplo 1.2.2.

1. *[Permutaciones con objetos diferentes.] Las formas posibles de sentar en forma lineal a las tres hermanas:* ***A****ndrea,* ***B****erenice y* ***C****armen, se representan por las*

siguientes seis permutaciones,

$$(ABC), (ACB), (BAC), (BCA), (CAB), (CBA).$$

2. *[Permutaciones con objetos repetidos.] Los números binarios de 3 dígitos, vienen dados por las siguientes 8 permutaciones.*

$$(000),\ (001),\ (010),\ (011),\ (100),\ (101),\ (110),\ (111).$$

Los cuales en el sistema decimal son: 0, 1, 2, 3, 4, 5, 6, 7, respectivamente.

3. *[Permutación de palabra.] Haremos referencia como permutaciones de una palabra a todas las permutaciones de los objetos de una permutación fija (la cual se puede considerar como la palabra).*
Por ejemplo las permutaciones de la palabra (mama) (ver Figura 1.1).

$$(aamm), (amam), (amma), (maam), (mama), (mmaa)$$

1.2.1. Diagrama de árbol

Cuando la cantidad de objetos a ordenar es pequeña, podemos hallar sus permutaciones construyendo un pequeño diagrama de árbol.
Por ejemplo hallemos las seis permutaciones de la palabra (mama).

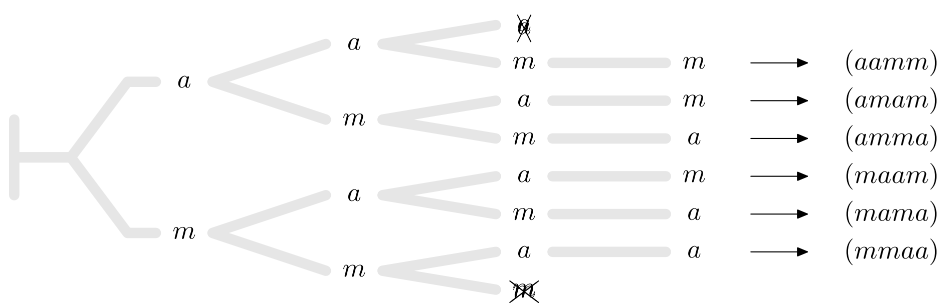


Figura 1.1: *Diagrama de árbol.*

Estas permutaciones se obtienen ordenando las letras de cada una de las **ramas con 4 letras**.

Definición 1.2.3. *La combinación de r objetos, es el arreglo $\{s_1, s_2, \ldots, s_r\}$ de dichos objetos sin tomar en cuenta su orden.*
Así, $\{s_1, s_2, \ldots, s_r\} = \{s_r, \ldots, s_2, s_1\}$ (son la misma combinación).

Ejemplo 1.2.4.

1. *[Combinaciones con objetos diferentes] Diagonales de un pentágono.*
   Una diagonal, es una recta que va de un vértice a otro no contiguo.
   La diagonal que va del vértice V_i al vértice V_j la cual es la misma diagonal que va de V_j al vértice V_i, la representaremos por la combinación $\{V_i, V_j\}$.
   Obsérvese (figura 1.2) que el pentágono tiene cinco diagonales: $d_1, \ldots, d_5$, representadas por las combinaciones respectivas:

$$\{V_1, V_4\}, \{V_2, V_5\}, \{V_3, V_1\}, \{V_4, V_2\}, \{V_5, V_3\}$$

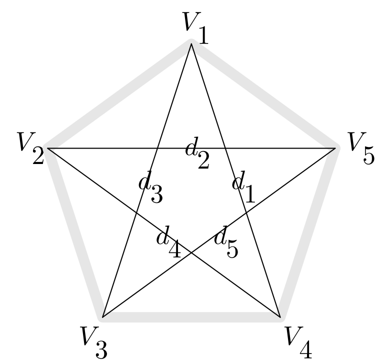


Figura 1.2: *Diagonales del pentágono.*

2. *[Combinaciones con objetos repetidos] Formas posibles de pedir un helado doble (2 bolas de helado) con sabores existentes: fresa, mamey y vainilla.*
   Si el helado de mamey y vainilla es representado por la combinación $\{m, v\}$, entonces siguiendo esta representación, las seis formas posibles de servir uno de estos helados son,

$$\{f, f\}, \{m, m\}, \{v, v\}, \{f, m\}, \{f, v\}, \{m, v\}$$

1.3. Permutaciones y Combinaciones sin repetición

Teorema 1.3.1. *Consideremos en total n objetos* ***diferentes****. La cantidad P_k^n de permutaciones de k de estos objetos tomados del total es:*

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

donde $r!$ conocido como r-factorial, representa al producto: $r! = 1(2)\cdots(r)$.

Demostración.

Etapas	1	2	$\cdots$	i	$\cdots$	k
No. de formas	n	$n-1$	$\cdots$	$n-(i-1)$	$\cdots$	$n-(k-1)$

Tarea: Construir las permutaciones (ordenaciones) de los k objetos.
Etapa 1: Tomar un objeto del total como primer elemento de la permutación. Esta etapa se puede realizar de n formas, ya que en total hay n objetos.
Para $i = 2, 3, \ldots, k$.
Etapa i: Tomar un objeto como i-ésimo elemento de la permutación, esta etapa se puede realizar de $n-(i-1)$ formas, ya que hemos utilizado (ordenado) $i-1$ objetos del total y nos quedan por ordenar $n-(i-1)$.
Luego por el principio de la multiplicación:

$$P_k^n = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

□

Corolario 1.3.2. *La cantidad P_n^n de permutaciones de n objetos **diferentes**, es,*

$$P_n^n = n!$$

Ejemplo 1.3.3. ***Ja**vier, **Je**sús, **Jo**sé y **Ju**an formaron una banda compuesta de 4 instrumentos **diferentes**.*
a) Si cada uno de los chicos puede utilizar los 4 instrumentos, ¿cuántos arreglos diferentes son posibles?
b) ¿Qué pasa si Javier y Juán pueden utilizar los 4 instrumentos, pero Jesús y José únicamente pueden utilizar los 2 instrumentos: piano y batería?

*a) Tarea: Asignar instrumento a cada músico: **Ja**vier, **Je**sús, **Jo**sé, **Ju**an .*

Etapas	*Ja*	*Je*	*Jo*	*Ju*
No. de formas	4	3	2	1

*b) En este inciso empezamos asignando instrumento a **Je**sús y a **Jo**sé los cuales sólo pueden utilizar 2 instrumentos (piano y batería).*

Etapas	*Je*	*Jo*	*Ja*	*Ju*
No. de formas	2	1	2	1

Tarea: Asignar instrumento a cada músico.
Etapa Je: Asignar instrumento a Jesús ya sea el piano o batería.
Etapa Jo: Asignar instrumento a José el instrumento que queda por asignar del piano y la batería.
Etapa Ja: Asignar cualquier de los 2 instrumentos restantes a Javier.
Etapa Ju: Asignar el último instrumento que queda a Juán.
Por el Principio de la Multiplicación, el número de arreglos es: (2)(1)(2)(1)=4

Teorema 1.3.4. *Consideremos, n objetos* ***diferentes*** *en total. La cantidad C_k^n de combinaciones de k objetos, tomados del total es:*

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

Esta cantidad también es denotada por la expresión $\binom{n}{k}$, conocida como coeficiente binomial.

Demostración. Obsérvese que cada combinación $\{s_1, \ldots, s_k\}$ de k objetos diferentes por Corolario 1.3.2 genera en total $k!$ permutaciones diferentes. Por ejemplo, la combinación $\{a, b, c\}$, genera las siguientes 3!=6 permutaciones, $(abc), (acb), (bac), (bca), (cab), (cba)$. Por lo que la relación entre las cantidades de Combinaciones y Permutaciones con k objetos diferentes viene dado por: $k!\ C_k^n = P_k^n$.
Por Teorema 1.3.1, $C_k^n = \dfrac{P_k^n}{k!} = \left(\dfrac{n!}{(n-k)!}\right)/k! = \dfrac{n!}{k!\ \cdot\ (n-k)!}$. □

Ejemplo 1.3.5. *A una reunión asisten 10 amigos y entre todos se intercambian saludos. ¿Cuántos saludos se han intercambiado?*
Los intercambios de saludos serán representados por combinaciones, así, $\{A_i, A_j\}$ indica el intercambio de saludo entre los amigos A_i, A_j, por lo que hay,

$$C_2^{10} = \binom{10}{2} = 45 \text{ intercambios de saludos.}$$

1.3.1. Permutaciones de Bloques ordenados

Corolario 1.3.6. *El total de permutaciones lineales de Bloques ordenados B_i de permutaciones con n_i objetos diferentes $i = 1, 2, \ldots, k$, viene dado por*

$$n_1! \cdot n_2! \cdots n_k! \cdot k!$$

Demostración.

Etapa	*elaborar bloque B_1*	$\cdots$	*elaborar bloque B_k*	*permutar los k bloques*
No. de formas	$n_1!$	$\cdots$	$n_k!$	$k!$

□

Ejemplo 1.3.7. *¿De cuántas maneras podemos sentar a tres niños y tres niñas en una fila:*

1. *sin restricción alguna?*

2. *si las tres hermanas Itzel , Nictehá y Rosa se sentarán juntas ?*

3. *si los niños como las niñas estarán sentados juntos?*

Los Bloques serán formados por individuos que se sentarán juntos.
1. Aquí se forman 6 Bloques de un elemento cada uno por lo que habrá 6!=720 formas de sentar a todos sin restricción.
2. Aquí formaremos 4 Bloques, los 3 primeros con un elemento formado con cada uno de los tres niños que no se sentarán juntos y el cuarto Bloque formado con las 3 hermans, por lo que habrá $1! \cdot 1! \cdot 1! \cdot 3! \cdot 4! = 144$ *formas de sentar a los seis, con la condición de que las niñas se sentarán juntas.*
3. Dos Bloques uno de niñas y el otro de niños con 3 elementos cada uno, por lo que habrá $3! \cdot 3! \cdot 2! = 72$ *formas de sentar a los seis, con la condición de que las niñas y los niños se sentarán juntos.*

1.3.2. Combinaciones de varios tipos

Teorema 1.3.8. *Consideremos* $n = n_1 + \cdots + n_r$ *objetos* ***diferentes*** *en total, donde* n_i *de ellos son del tipo* i*, para* $i = 1, \ldots, r$*. La cantidad de combinaciones de* $k = k_1 + \cdots + k_r$ *objetos tomados del total, donde* k_i *de ellos son del tipo* i*, para* $i = 1, \ldots, r$*, viene dada por:*

$$\binom{n_1}{k_1} \cdot \binom{n_2}{k_2} \cdots \binom{n_r}{k_r}$$

Demostración.

$Etapas$	1	$\cdots$	i	$\cdots$	r
$N_0\ de formas$	$\binom{n_1}{k_1}$	$\cdots$	$\binom{n_i}{k_i}$	$\cdots$	$\binom{n_r}{k_r}$

Tarea: Construcción de las combinaciones en r etapas. Para $i = 1, \ldots r$:
Etapa i: Construcción de combinaciones de tipo i con k_i objetos, tomados del total con n_i del tipo i. Por el Teorema1.3.4, esta etapa se realiza de $\binom{n_i}{k_i}$ formas.
Por el principio de la multiplicación, la cantidad buscada de combinaciones es:

$$\binom{n_1}{k_1} \cdot \binom{n_2}{k_2} \cdots \binom{n_r}{k_r}$$

□

Ejemplo 1.3.9. *Un estudiante recibirá de obsequio 2 libros, de una colección de: 7 libros de matemáticas, 6 de ciencias y 4 de economía. ¿De cuántas maneras se puede hacer, si:*

1. *Ambos libros serán del mismo tema? (Combinaciones de un solo tipo)*

$$C_2^7 + C_2^6 + C_2^4 = \binom{7}{2} + \binom{6}{2} + \binom{4}{2} = 42.$$

2. *Los libros serán sobre temas diferentes? (Combinaciones de dos tipos)*

$$C_1^7\ C_1^6 + C_1^7\ C_1^4 + C_1^6\ C_1^4 = \binom{7}{1}\binom{6}{1} + \binom{7}{1}\binom{4}{1} + \binom{6}{1}\binom{4}{1} = 94.$$

1.4. Permutaciones y combinaciones con repetición

1.4.1. Permutación de palabras

Ya en el elemplo 1.2.2 se habló de permutación de una palabra.

Teorema 1.4.1. *La cantidad de* ***permutaciones de una palabra*** *con $n = n_A + n_B + \cdots + n_Z$ letras en total, donde n_α de ellas son letras α (para $\alpha = A, \ldots Z$), viene dada por,*

$$\binom{n}{n_A, \ldots, n_z} = \frac{n!}{n_A!\ \cdot\ n_B!\ \cdot\ n_C!\ \cdots\ n_Z!}$$

La expresión $\binom{n}{n_1,\ldots,n_r}$ es conocida como coeficiente multinomial.

Demostración.

Etapas	A	B	$\cdots$	α	$\cdots$	Z
N_0 *de formas*	$\binom{n}{n_A}$	$\binom{n-n_A}{n_B}$	$\cdots$	$\binom{n-n_A-\cdots-n_{\alpha-1}}{n_\alpha}$	$\cdots$	$\binom{n-n_A-\cdots-n_Y}{n_Z}$

Sin pérdida de generalidad podemos suponer que $n_\alpha \geq 1$, para $\alpha = A, \ldots Z$.
Tarea: Escribir las letras de dicha palabra en las etapas $A, B, \ldots, Z$.
Etapa A: Tomar n_A de estos n lugares, para escirbir en ellos, la letra A, por Teorema 1.3.4, esta etapa se realiza de $\binom{n}{n_A}$ formas.
Etapa B: Tomar n_B lugares de los $n - n_A$ restantes, para escirbir en ellos la letra B, por Teorema 1.3.4, esta etapa se realiza de $\binom{n-n_A}{n_B}$ formas.
Para $\alpha = C, \ldots, Z$ y siguiendo el orden Lexicográfico: $A, B, C, \ldots, Z$.
Etapa α: Tomar n_α lugares de los $n - n_A - \cdots - n_{\alpha-1}$ restantes, para escirbir en ellos la letra α, por Teorema 1.3.4, esta etapa se realiza de $\binom{n-n_A-\cdots-n_{\alpha-1}}{n_\alpha}$ formas.
Como $(n - n_A - \cdots - n_Y - n_Z)! = 0! = 1$, por el principio de la multiplicación la

cantidad de permutaciones de dicha palabra es,

$$
\begin{aligned}
&\binom{n}{n_A}\binom{n-n_A}{n_B}\cdots\binom{n-n_A-\cdots-n_{\alpha-1}}{n_\alpha}\cdots\binom{n-n_A-\cdots-n_Y}{n_Z}\\
=&\frac{n!}{(n_A)!\ \cancel{(n-n_A)!}}\frac{\cancel{(n-n_A)!}}{(n_B)!\ \cancel{(n-n_A-n_B)!}}\cdots\frac{\cancel{(n-n_A-\cdots-n_Y)!}}{(n_Z)!\ (n-n_A-\cdots-n_Y-n_Z)!}\\
=&\frac{n!}{n_A!\ \cdot\ n_B!\ \cdot\ n_C!\ \cdots\ n_Z!}
\end{aligned}
$$

□

Ejemplo 1.4.2. *Hallemos la cantidad de permutaciones de la palabra mama, tal cantidad se obtiene con la fórmula anterior y las permutaciones se obtubieron con el diagrama de árbol (ver Figura 1.1) .*

$$\frac{n!}{n_A!\ \cdot\ n_M!}=\frac{4!}{2!\ \cdot\ 2!}=6$$

las 6 permutaciones son: (aamm),(amam),(amma),(maam),(mama),(mmaa).

Ejemplo 1.4.3. *Un mecanismo puede ponerse en cuatro posiciones, digamos a, b, c y d. Hay ocho de tales mecanismos en un Sistema (ver Figura 1.3). Suponga que dichos mecanismos están instalados en algún orden lineal preasignado.*

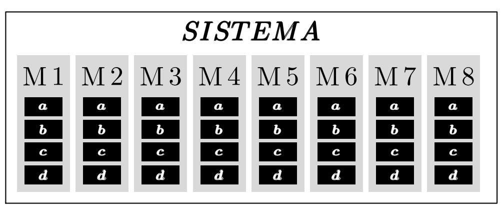


Figura 1.3: *Sistema.*

a) *¿De cuántas maneras puede instalarse el sistema?*

b) *¿De cuántas maneras posibles se instalan los mecanismos, si dos mecanismos adyacentes no están en la misma posición?*

c) *¿Cuántas maneras son posibles si solo se usan las posiciones a y b con la misma frecuencua?*

d) *¿Cuántas maneras son posibles si solo se usan dos posiciones diferentes y una de ellas aparece tres veces más que la otra?*

Con permutaciones describiremos las maneras de isntalar los mecanismos del sistema, así, (acccccccd) indica que el primer mecanismo está en la posición a, el octavo en la posición d y el resto de los mecanismos están en la posición c.
a) Ejemplos: $(a, a, a, a, a, a, a, a), (d, d, d, d, d, d, d, d)$

$Etapas$	1	2	$\cdots$	8
N_0 de $formas$	4	4	$\cdots$	4

Tarea: Instalar el sistema que consta de 8 mecanismos.
Para $i = 1, \ldots, 8$,
Etapa i: instalar el i-ésimo mecanismo, el cual se puede poner en cuatro posiciones.
Por Principio de la Multiplicación hay $4^8 = 65\ 536$ *maneras de instalar el sistema.*
b) Ejemplos: $(a, b, c, d, a, b, c, d), (a, b, a, b, a, b, a, b)$.

$Etapas$	1	2	$\cdots$	8
N_0 de $formas$	4	3	$\cdots$	3

Tarea: Instalar el sistema, sin mecanismos adyacentes en la misma posición.
Etapa 1*: instalar el primer mecanismo, el cual se puede poner en cuatro posiciones.*
Para $i = 2, \ldots, 8$,
Etapa i: instalar el i-ésimo mecanismo, el cual se puede poner en tres posiciones, ya que la posición anterior no se puede utilizar.
Por Principio de la Multiplicación hay $4(3^7) = 8\ 748$ *maneras de instalar el sistema.*
c) Ejemplos: $(a, a, a, a, b, b, b, b), (b, b, b, b, a, a, a, a)$.
En este inciso no utilizaremos el Principio de la Multiplicación, porque se presentan permutaciones de una palabra con 4 letras a y 4 letras b, luego habrá,

$$\frac{n!}{n_a! \cdot n_b!} = \frac{8!}{4! \cdot 4!} = 70$$

maneras de instalar el sistema usando las posiciones a y b con la misma frecuencia.
d) Ejemplos: $(a, a, b, b, b, b, b, b), (b, b, a, a, a, a, a, a)$.

$Etapas$	1	2	3
N_0 de $formas$	4	3	28

Tarea: Construir en tres etapas: palabras con 2 posiciones diferentes donde una se repita 2 veces y la otra 6.
Etapa 1: Seleccionar una primera letra del conjunto $\{a, b, c, d\}$ *la cual aparecerá 2 veces en la palabra.*
Etapa 2: Seleccionar una segunda letra del conjunto $\{a, b, c, d\}$ *que sea diferente de la primera letra seleccionada anteriormente la cual aparecerá 6 veces en la palabra.*
Etapa 3: Con las 2 letras seleccionadas en las 2 etapas anteriores, construir las permutaciones de la palabra correspondiente. Esta etapa se realiza de $\dfrac{8!}{6! \cdot 2!} = 28$ *formas.*
Por Principio de la Multiplicación hay 4(3)(28)=336 maneras de instalar el sistema, usando dos posiciones diferentes donde una de ellas aparece tres veces más que la otra.

Nota 1.4.4 (Estrategia al trabajar permutaciones). *Mi estrategia cuando utilizo permutaciones es que si no son **permutaciones de una sola palabra**, entonces utilizo el Principio de la multiplicación. **Observa** el uso de la estrategia en el ejemplo anterior.*

1.4.2. Combinaciones con repetición

Nota 1.4.5. *1. Desarrollando la expresión algebraica $(x_1+x_2+\cdots+x_r)^n$, obtenemos como términos a todas las palabras con n letras en total conteniendo al menos una de las r letras $x_1, x_2, \ldots, x_r$.*

*2. Sumando términos semejantes, obtenemos el **Teorema multinomial**,*

$$(x_1+x_2+\cdots+x_r)^n = \sum_{\substack{(n_1,\ldots,n_r)\in(\mathbb{N}^+)^n:\\ n_1+\cdots+n_r=n\\ \mathbb{N}^+=\mathbb{N}\cup\{0\}}} \binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1}\ x_2^{n_2}\cdots x_r^{n_r}$$

donde $\binom{n}{n_1,n_2,\ldots,n_r} x_1^{n_1}\ x_2^{n_2}\cdots x_r^{n_r}$, es la suma de los términos semejantes que tienen como parte literal $x_1^{n_1}\ x_2^{n_2}\cdots x_r^{n_r}$, donde, $n_1+\cdots+n_r=n$,
$n_i \in \mathbb{N}^+,\ i=1,2,\ldots r.$.
Por lo tanto, habrá tantas combinaciones con repetición de tamaño n de al menos uno, de los r objetos $x_i, \ldots, x_r$; como términos tenga la sumatoria en el Teorema multinomial.

Definición 1.4.6. *Una r-partición entera del número natural N, es una permutación con repetición de r números naturales $N_1, N_2 \ldots, N_r$, tales que:*

$$N_1+N_2+\cdots+N_r = N \qquad (*)$$

Obsérvese que una r-partición entera del número natural N no es más que una ***r*-solución de la ecuación (*) en el conjunto de los números naturales $\mathbb{N}$**.

Ejemplo 1.4.7. *Las 3-particiones enteras del número natural 5, son:*

$$(1,1,3); (1,2,2); (1,3,1); (2,1,2); (2,2,1); (3,1,1)$$

El siguiente Teorema nos dice que hay 6 de estas 3-particiones en total.

Teorema 1.4.8. *La cantidad de r-particiones enteras del número natural N, equivalentemente, la cantidad de r-soluciones de la ecuación (*) en el conjunto de los números naturales $\mathbb{N}$, viene dada por:*

$$\binom{N-1}{r-1}$$

Demostración. La demostración se sigue de la siguiente afirmación y del Teorema 1.3.4. Es un buen ejercicio demostrar la afirmación: La función H es una biyección entre el conjunto de combinaciones de r números naturales menores que N, y el conjunto de r-particiones enteras del número natural N,

$$H(\{k_1, k_2, \ldots, k_{r-1}\}) = (k_1, k_2 - k_1, \ldots, N - k_{r-1})$$

donde $0 < k_1 < k_2 < \ldots < k_{r-1} < N$ (ver Figura 1.4), □

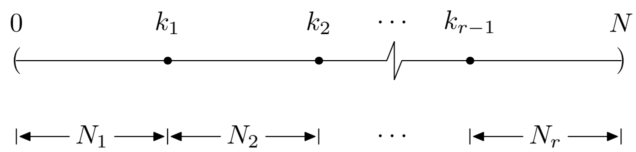


Figura 1.4: *Partición de un número entero.*

Teorema 1.4.9. *La cantidad de combinaciones de tamaño n con repetición de al menos uno de r objetos diferentes, es denotada por CR_r^n y viene dada por,*

$$CR_r^n = \binom{n+r-1}{r-1}$$

Demostración. Por Nota 1.4.5 CR_r^n viene dada por el número de términos que tiene la suma del Teorema multinomial, el mismo Teorema nos dice, que esta cantidad CR_r^n es la cardinalidad del conjunto A de r-soluciones en los **enteros no negativos** de la ecuación:

$$n_1 + \cdots + n_r = n \Longleftrightarrow (n_1 + 1) \cdots + (n_r + 1) = n + r,$$

haciendo $0 < n_i + 1 = N_i \in \mathbb{N},\ i = 1, \ldots, r$, obtenemos la ecuación,

$$N_1 + \cdots + N_r = n + r \qquad (2)$$

Por la biyección $G(n_1, n_2, \ldots, n_r) = (N_1, N_2, \ldots, N_r)$ entre el conjunto A y el conjunto B de r-soluciones en los **números naturales $\boldsymbol{N}$** de la ecuación (2), y por el Teorema 1.4.8 aplicado en la ecuación (2), se obtiene que:

$$CR_r^n = \binom{n+r-1}{r-1}$$

□

Ejemplo 1.4.10. *En un edificio de 6 pisos, un ascensor arranca en el sótano con 8 personas. (sin incluir al operador del ascensor). ¿De cuántas maneras podría el operador haber dejado a las personas que salen del ascensor, si el operador no puede distinguir entre ellos?*

Cualquier resultado observado por el operador lo podemos expresar de las dos siguientes formas equivalentes,

polinomio		*combinación*
N_6^8	$\equiv$	$\{N_6N_6N_6N_6N_6N_6N_6N_6\}$
$N_1^3N_2^1N_3^1N_4^1N_5^1N_6^1$	$\equiv$	$\{N_1N_1N_1N_2N_3N_4N_5N_6\}$

donde en el primer ejemplo se indica que las 8 personas bajaron del ascensor en el último nivel, y en el segundo ejemplo se expresa que 3 personas bajaron del ascensor en el nivel 1 y en los otros cinco niveles bajó solamente 1 persona. Obsérvese que nos piden, la cantidad de combinaciones de tamaño $n = 8$, de al menos uno de los $r = 6$ objetos diferentes: $N_1, N_2, N_3, N_4, N_5, N_6$.
Por el Teorema anterior,

$$CR_6^8 = \binom{n+r-1}{r-1} = \binom{13}{5} = 1\ 287.$$

Ejemplo 1.4.11. *Regresando al Ejemplo 1.2.4 ¿De cuántas formas posibles se puede pedir un helado triple (3 bolas de helado), si los sabores existentes son: fresa, mamey y vainilla?*
Las solicitudes tienen la forma de combinaciones con repetición $\{s_1, s_2, s_3\}$ de tamaño $n = 3$, de los $r = 3$ sabores s_i: fresa, mamey y vainilla.
Luego la cantidad de formas de servir los helados es,

$$CR_3^3 = \binom{n+r-1}{r-1} = \binom{5}{2} = 10$$

Las 10 formas de servir los helados son: $\{f, f, f\}; \{m, m, m\}; \{v, v, v\}; \{f, f, m\}; \{f, f, v\}; \{m, m, f\}; \{m, m, v\}; \{v, v, f\}; \{v, v, m\}; \{f, m, v\}$.

1.5. Ejercicios

1. Entre un grupo de 15 personas se debe elegir una mesa directiva que conste de un presidente, un secretario y dos vocales. ¿De cuántas maneras se puede hacer esta selección?

2. De un grupo de 10 personas debe elegirse un comité de cuatro miembros. Pero, por ciertas razones, en él debe figurar forzosamente Don Perpetuo o su hermano, pero no los dos, pues ‟se vería mal˝. ¿Cuántos comités diferentes se pueden formar?

3. Una baraja completa consta de 52 cartas (13 números:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K; y cuatro palos: $\heartsuit, \diamondsuit, \clubsuit, \clubsuit$). Una mano de póker consta de cinco cartas. ¿Cuántas manos de póker habrá con todos los números diferentes? Sugerencia, Tarea realizada en 2 etapas: Etapa1, poner número a mano de 5 cartas; Etapa 2, poner palo a las cinco cartas ya con número de la etapa anterior.

4. En el dominó hay 28 fichas, de las cuales 7 son dobles. Una mano consta de 7 fichas. ¿Cuántas manos de dominó hay que tengan exactamente cuatro fichas dobles?

5. ¿Cuántos números naturales hay de cuatro dígitos, que sean múltiplos de cinco, menores que 7 000, y mayores que 4 000?

6. ¿Cuántos números naturales de tres dígitos no contienen dígitos repetidos.

7. En una clase hay 10 alumnos y se repartirán 3 premios ¿De cuántas formas diferentes se puede realizar si:

   a) los premios son iguales?

   b) los premios son diferentes?

   Considere los casos cuando la persona puede recibir más de un premio, o solamente uno.

8. Una línea de ferrocarril tiene 25 estaciones ¿Cuántos boletos diferentes habrá que imprimir, si cada boleto lleva impresas las estaciones de origen y destino?

9. ¿De cuántas maneras se puede sentar a 5 hombres con 4 mujeres de modo que las mujeres ocupen los lugares pares?

10. ¿Cuántos números naturales hay de cuatro dígitos:

    a) con repetición?

    b) sin repetición?

    c) sin repetición, que sean pares?

    d) con repetición, que sean múltiplos de cinco?

11. Al lanzar 4 monedas ¿Cuántos resultados posibles se pueden obtener, y en cuántos de éstos hay dos águilas con dos soles?

12. Con las letras de la palabra libro, ¿cuántas ordenaciones distintas se pueden hacer que empiecen por vocal?

13. ¿Cuántos números impares se pueden formar con cinco dígitos distintos, y cuántos de éstos son mayores que 70 000?

14. ¿De cuántos partidos consta una liguilla formada por cuatro equipos? Tómese en cuenta que no es lo mismo ser visitante que ser local.

15. A una reunión asisten 10 personas y se intercambian saludos entre todos. ¿Cuántos saludos se han intercambiado?

16. Con las cifras 1, 2 y 3, ¿cuántos números de cinco cifras pueden formarse, y cuántos son pares?

17. ¿De cuántas formas pueden colocarse en un partido a 11 jugadores de un equipo de fútbol, teniendo en cuenta que el portero no puede ocupar otra posición distinta de la portería?

18. Con punto y raya como pulsaciones en el Sistema Morse, ¿cuántas señales distintas se pueden enviar, usando cuatro o menos pulsaciones en cada señal?

19. ¿Cuántas diagonales tiene un pentágono, y cuántos triángulos se pueden formar con sus vértices?

20. Cuatro libros de matemáticas, seis de física y dos de química han de ser colocados en una estantería. Si los libros de cada uno de los tres bloques son iguales, ¿cuántas ordenaciones distintas admite si:

    a) los libros de cada materia han de estar juntos?

    b) sólo los de matemáticas han de estar juntos?

21. Un alumno tiene que elegir 7 de las 10 preguntas de un examen.

    a) ¿De cuántas maneras puede elegirlas?

    b) ¿De cuántas si las cuatro primeras son obligatorias?

22. De un grupo compuesto por cinco hombres y siete mujeres se eligen 2 hombres y 3 mujeres para formar un comité . ¿De cuántas formas puede formarse si:

    a) puede pertenecer a él cualquier hombre o mujer?

    b) una mujer determinada debe pertenecer al comité?

    c) dos hombres determinados no pueden estar en el comité?

23. Durante años, los códigos de área telefónicos en los Estados Unidos y Canadá consistieron en una secuencia de tres dígitos. El primer dígito era un número entero entre 2 y 9, el segundo dígito era 0 ó 1, y el tercer dígito era cualquier número entero del 1 al 9. ¿Cuántos códigos de área fueron posibles y cuántos de éstos comienzan con dígito impar?

24. ¿De cuántas maneras pueden sentarse linealmente tres niños y tres niñas si:

    a) los niños como las niñas estarán sentados juntos?

    b) solamente los niños deberán sentarse juntos?

    c) no habrá dos personas del mismo sexo sentadas juntas?

25. Un niño tiene 12 bloques iguales, de los cuales 6 son negros, 4 son rojos, 1 es blanco, y 1 es azul. Si el niño pone los bloques en una línea, ¿cuántos arreglos son posibles?

26. ¿De cuántas maneras se pueden sentar 8 personas en fila si:

    a) no hay restricciones en los asientos.
    b) Andrea y Berenice deben sentarse juntas?
    c) hay 4 hombres, y 2 hombres no pueden sentarse juntos?
    d) hay 5 hombres y éllos deben sentarse juntos?
    e) hay 4 parejas casadas y los esposos deben sentarse juntos?

27. ¿De cuántas maneras, 3 novelas, 2 libros de matemáticas, y 1 libro de química se pueden colocar en orden lineal si:

    a) los libros pueden ser ordenados en cualquier orden?
    b) los libros de matemáticas deben estar juntos y las novelas también?
    c) las novelas deben estar juntas y el resto en cualquier orden?

28. Cinco premios separados (mejor beca, mejor cualidades de pertenencia, y así sucesivamente) deben ser otorgados a estudiantes seleccionados de una clase de 30. ¿Cuántos resultados son posibles si:

    a) un estudiante puede recibir cualquier cantidad de premios?
    b) un estudiante no puede recibir más de un premio?

29. Una clase de danza consta de 22 estudiantes, de los cuales 10 son mujeres y 12 son hombres. ¿Cuántos grupos diferentes se pueden formar:

    a) con 10 integrantes en los grupos?
    b) con 5 mujeres y 5 hombres?
    c) con 5 parejas (1 pareja = 1 hombre y 1 mujer)?

30. Siete regalos diferentes serán distribuidos entre 10 niños. ¿Cuántos resultados distintos son posibles Si ningún niño va a recibir más de un regalo?

31. Un comité de 7, compuesto por 2 republicanos, 2 demócratas, y 3 independientes, serán elegidos respectivamente de un grupo de 5 republicanos, 6 demócratas, y 4 independientes. ¿Cuántos comités serán posibles?

32. De un grupo de 8 mujeres y 6 hombres, un comité consistiendo de 3 hombres y 3 mujeres será formado. ¿Cuántos comités diferentes son posibles si:

    a) 2 de los hombres se niegan a servir juntos?

b) 2 de las mujeres se niegan a servir juntas?

c) 1 hombre y 1 mujer se niegan a servir juntos?

33. Una persona tiene 8 amigos, de los cuales 5 serán invitados a una fiesta.

    a) ¿Cuántas opciones hay si 2 de los amigos están peleados y no van a asistir juntos?

    b) ¿Cuántas opciones, si 2 de los amigos, solo asistirían, si van juntos?

34. Un laboratorio de psicología que realiza investigaciones sobre los sueños contiene 3 habitaciones, con 2 camas en cada habitación. Si se deben asignar 3 parejas de hermanos a estas 6 camas, para que cada pareja de hermanos duerman en camas diferentes en la misma habitación, ¿cuántas asignaciones son posibles?

35. Si 12 personas se dividen en 3 comités de tamaños respectivos 3, 4 y 5, ¿cuántas divisiones son posibles?

36. Si 8 nuevos profesores se reparten entre 4 escuelas, ¿cuántas divisiones son posibles? Y ¿cuántas si cada escuela debe recibir 2 profesores?

37. Diez levantadores de pesas compiten en un concurso de levantamiento de pesas en equipo. De los cuales, 3 son de Estados Unidos, 4 de Rusia, 2 de China, y 1 de Canada. Si el puntaje tiene en cuenta al país que representa al levantador, pero no sus identidades individuales, ¿cuántos resultados diferentes son posibles desde el punto de vista de los puntajes? ¿En cuántos de estos resultados, los Estados Unidos tiene 1 competidor entre los tres primeros lugares y 2 en los tres últimos lugares?

38. Los delegados de 10 países, incluidos Rusia, Francia, Inglaterra y los Estados Unidos, se sentarán en una fila. ¿Cuántos arreglos diferentes son posibles si los delegados franceses e ingleses deben sentarse uno al lado del otro y los delegados de Rusia y Estados Unidos no deben estar uno junto al otro?

39. Si 8 pizarras idénticas se reparten entre 4 escuelas.

    a) ¿Cuántas reparticiones son posibles?

    b) ¿Cuántas, si cada escuela debe recibir al menos una pizarra?

40. En un edificio de 6 pisos, un ascensor arranca en el sótano con 8 personas. (sin incluir al operador del ascensor). ¿De cuántas maneras podría el operador haber dejado a las personas que salen del ascensor, si:

    a) El operador no puede distinguir entre ellos?

    b) En las 8 personas hay 5 hombres y 3 mujeres y el operador solo puede distinguir si es hombre ó es mujer?

c) El operador no puede distinguir entre ellos y el ascensor se descarga por completo en el piso número 6 ?

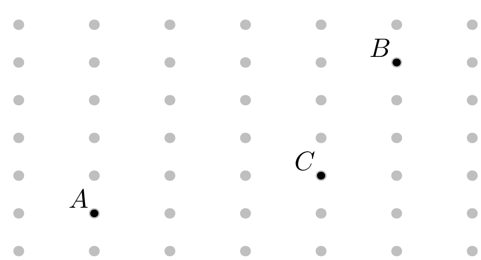
41. Tenemos 20 mil dólares que se van a invertir entre 4 posibles oportunidades. Cada inversión debe ser integral en unidades de 1 mil dólares, y hay inversiones mínimas que necesitan ser hechas en estas oportunidades. Las inversiones mínimas en las oportunidades A, B, C y D son 2, 2, 3 y 4 mil dólares respectivamente.
¿Cuántas estrategias de inversión diferentes están disponibles si:

    a) Deberá hacerse una inversión en cada oportunidad?
    b) Las inversiones deben realizarse en al menos 3 de los 4 oportunidades?

42. Una colección de arte en subasta consistió en 4 Dalis, 5 Van Goghs, y 6 de Picassos. En la subasta se presentaron 5 coleccionistas de arte. Si un reportero anotó sólo el número de Dalis, van Goghs y Picassos adquiridos por cada uno de los coleccionistas, ¿cuántos resultados diferentes podría tener registrados, si todas las obras fueron vendidas?

43. Considera la siguiente cuadrícula de puntos. Supongamos que, comenzando en el punto A, se puede ir un paso hacia arriba o un paso hacia la derecha en cada movimiento. Este procedimiento continúa hasta:

    a) que el punto B es alcanzado. ¿Cuántos caminos diferentes son posibles?
    b) que el punto B es alcanzado, pero pasando por el punto C. ¿Cuántos caminos diferentes son posibles?



Capítulo 2

Conceptos y Teoremas básicos de probabilidad

La probabilidad es el lenguaje de la estadística. Es este capítulo empezamos su construcción con el estudio de los conceptos y teoremas básicos de esta teoría rica en aplicaciones en una amplia variedad de problemas en la ingeniería y la ciencia (ver [2], [3]).

2.1. Elementos básicos

2.1.1. Experimento aleatorio y Espacio muestral

Definición 2.1.1.

1. *Un experimento aleatorio* ϵ*, es aquel que tiene al menos dos posibles resultados.*
2. *El espacio muestral* Ω *del experimento aleatorio* ϵ*, es el conjunto de todos los posibles resultados de dicho experimento.*

Ejemplo 2.1.2.

1. *El experimento más simple* ϵ*, consiste en lanzar una moneda y observar su cara superior, en este caso, hay dos posibles resultados: águila* (A)*, o sol* (S)*, por lo que* $\Omega = \{ A, S \}$.
2. *Un experimento con seis posibles resultados lo obtenemos al lanzar un dado y observar la cara superior, en este caso,* $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$.
3. *En general el experimento* ϵ*, que consiste en extraer al azar una esfera de una urna conteniendo* n *esferas etiquetadas con todos los números* $1, 2, \ldots, k$*, (con* $n \geq k$*), tiene como espacio muestral a:*

$$\Omega = \{ 1, 2, \ldots, k \}.$$

Nota 2.1.3. *En probabilidad se trabaja con cierta colección $\mathscr{F}$ de subconjuntos $A \subset \Omega$ llamados eventos, a los cuales se les puede asignar un valor $P(A)$, el cual es conocido como la probabilidad de que ocurra el evento A, que nos mide el grado de confiabilidad de que ocurra el evento A, dicha colección es conocida como σ álgebra.*

2.1.2. Espacio de Probabilidad

Definición 2.1.4. *Un espacio de probabilidad, es un triplete $(P, \Omega, \mathscr{F})$ donde:*

1. *Ω es un espacio muestral.*
2. *$\mathscr{F}$ es una σ-álgebra sobre Ω, es decir, $\mathscr{F}$ es una colección de subconjuntos de Ω, que satisface los siguientes tres axiomas*

   S_1 $\quad \Omega \in \mathscr{F}$.

   S_2 $\quad$ *Si $A \in \mathscr{F}$, entonces $A^c \in \mathscr{F}$.*

   S_3 $\quad$ *Si $A_1, A_2, \cdots \in \mathscr{F}$, entonces $\cup_1^\infty A_i \in \mathscr{F}$.*
3. *P es una probabilidad sobre $\mathscr{F}$, es decir:*

$$\begin{aligned} P : \mathscr{F} &\longrightarrow \mathbb{R} \\ A &\longmapsto P(A) \end{aligned}$$

   es una función que satisface los siguientes tres axiomas

   A_1 $\quad$ *Para toda $A \in \mathscr{F}$, $0 \leq P(A) \leq 1$.*

   A_2 $\quad P(\Omega) = 1$.

   A_3 $\quad$ *Para toda $A_1, A_2, \cdots \in \mathscr{F}$ de conjuntos mutuamente excluyentes o ajenos o disjuntos (es decir $A_i \cap A_j = \varnothing$ si $i \neq j$), se sigue*

$$P(\cup_{i=1}^\infty A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Teorema 2.1.5. *Sea $(P, \Omega, \mathscr{F})$ un espacio de probabilidad, y $A, B \in \mathscr{F}$, entonces:*

1. *$A = (A \cap B) \cup (A \cap B^c)$*
2. *$P(A) = P(A \cap B) + P(A \cap B^c)$.*

Demostración.

$$\begin{aligned} & A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c) \\ \text{además} \quad & (A \cap B) \cap (A \cap B^c) = A \cap (B \cap B^c) = A \cap \varnothing = \varnothing \end{aligned}$$

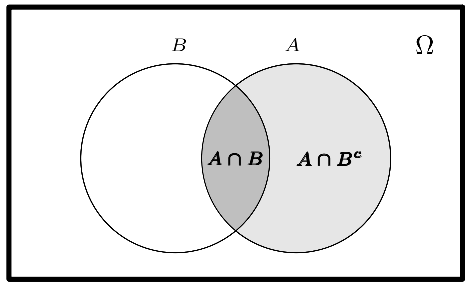


Figura 2.1: $A = (A \cap B) \cup (A \cap B^c)$

Siendo A unión de estos conjuntos disjuntos, por axioma A_3

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

$\square$

Teorema 2.1.6. *Sea $(P, \Omega, \mathscr{F})$ un espacio de probabilidad, y $A, B \in \mathscr{F}$, entonces:*

1. $P(\varnothing) = 0,$
2. $P(A^c) = 1 - P(A).$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B).$
4. $P(B - A) = P(B) - P(A \cap B).$
   Si además $A \subset B$, entonces:
   a) $P(B - A) = P(B) - P(A),$
   b) $P(A) \leq P(B).$

Demostración.

1. Por Teorema 2.1.5, $\Omega = (\Omega \cap \varnothing) \cup (\Omega \cap \varnothing^c) = \varnothing \cup \Omega, \quad P(\Omega) = P(\varnothing) + P(\Omega)$

$$\therefore\ P(\varnothing) = 1 - 1 = 0, \quad \text{(Por axioma } A_2\text{).}$$

2. Por Teorema 2.1.5, $\Omega = (\Omega \cap A) \cup (\Omega \cap A^c) = A \cup A^c, \quad P(\Omega) = P(A) + P(A^c)$

$$\therefore\ P(A^c) = 1 - P(A), \quad \text{(Por axioma } A_2\text{).}$$

3. Por Teorema 2.1.5, $A \cup B = ((A \cup B) \cap A) \cup ((A \cup B) \cap A^c) = A \cup (B \cap A^c),$

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

Por Teorema 2.1.5, $B = (B \cap A) \cup (B \cap A^c)$,

$$P(B) = P(A \cap B) + P(B \cap A^c) \qquad (2),$$

finalmente despejando $P(B \cap A^c)$ de (2), y sustituyendolo en (1), se obtiene

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

4. Por definición $B - A = B \cap A^c$, luego despejando $P(B \cap A^c)$ de (2), se obtiene:

$$P(B - A) = P(B) - P(A \cap B) \qquad (3).$$

Si además $A \subset B$, entonces $A \cap B = A$. luego de (3):

a) $P(B - A) = P(B) - P(A)$,

b) $0 \leq P(B - A) = P(B) - P(A)$, entonces $P(A) \leq P(B)$.

□

El siguiente Teorema que se demuestra por Inducción matemática, generaliza la afirmación 3. del Teorema 2.1.6

Teorema 2.1.7. *Sea* $(P, \Omega, \mathscr{F})$ *un espacio de probabilidad, y* $A_1, A_2, \ldots, A_n \in \mathscr{F}$, *entonces para* $n \geq 2$

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \cdots + \\ + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

Obsérvese por ejemplo que la tercera sumatoria, se realiza sobre el conjunto $\{(i, j, k) \mid 1 \leq i < j < k \leq n\} = \{\textit{Combinaciones } \{i, j, k\} \mid i, j, k \in \{1, 2, \ldots, n\}\}$, *por lo que esta sumatoria tendrá* $\binom{n}{3}$ *términos.*

2.1.3. Espacio Muestral Finito

Nota 2.1.8. *Cuando* Ω *es un conjunto finito o numerable, la* σ*-álgebra de* Ω *es el Conjunto Potencia* $\mathscr{P}(\Omega)$ *de* Ω, *es decir*

$$\mathscr{F} = \mathscr{P}(\Omega) = \{\ A \mid A \subset \Omega\ \}$$

Teorema 2.1.9. *Sean* Ω *espacio muestral finito, y* $A = \{a_1, \ldots, a_n\} \subset \Omega$.
Por axioma A_3 : $\quad P(A) = P(\{a_1\}) + P(\{a_2\}) + \cdots + P(\{a_n\})$.
Si además Ω *es equiprobable, es decir, todos los elementos de* Ω *tienen la misma probabilidad de ocurrencia, con* $p = P(\{w_i\})$ *para toda* $w_i \in \Omega$, *entonces*

$$P(A) = \frac{\#A}{\#\Omega},$$

donde $\#A$ *denota la cantidad de elementos del evento* A.

Demostración. Por axiomas A_1 y A_3

$$1 = P(\Omega) = \sum_{w_i \in \Omega} P(w_i) = \sum_{w_i \in \Omega} p = p\ (\#\Omega) \qquad \therefore \qquad p = \frac{1}{\#\Omega}$$

Análogamente

$$P(A) = \sum_{w_i \in A} P(w_i) = \sum_{w_i \in A} p = p\ (\#A) = \frac{1}{\#\Omega}(\#A)$$

$\square$

Ejemplo 2.1.10. *Entre los números 1,2,...,1000 se escoge uno al azar. ¿Cuál es la probabilidad de que el número escogido sea divisible entre 6 o entre 8?*
En este caso, el espacio muestral es

$$\Omega = \{1, 2, \ldots, 1000\}$$

Sea A_n =*{subconjunto de* Ω *cuyos elementos son números divisibles entre n}.*

Nos piden hallar $P(A_6 \cup A_8)$*, por Teorema 2.1.6,3.*

$$\begin{aligned}
\#A_6 = 166, &\quad \textit{ya que}, \quad \frac{1000}{6} = 166\ \tfrac{2}{3}, \ \textit{y}\ A_6 = \{\mathbf{1}(6), \mathbf{2}(6), \ldots, \mathbf{166}(6)\} \\
\#A_8 = 125, &\quad \textit{ya que}, \quad \frac{1000}{8} = 125, \\
\#(A_6 \cap A_8) = A_{24} = 41, &\quad \textit{ya que}, \quad mcm(6,8) = 24 \quad y \quad \frac{1000}{24} = 41\ \tfrac{2}{3}. \\
P(A_6 \cup A_8) = &\ P(A_6) + P(A_8) - P(A_6 \cap A_8) = \frac{\#A_6}{\#\Omega} + \frac{\#A_8}{\#\Omega} - \frac{\#(A_6 \cap A_8)}{\#\Omega} \\
= &\ \frac{166 + 125 - 41}{1000} = \frac{1}{4}
\end{aligned}$$

Ejemplo 2.1.11. *Al escribir linealmente y de manera aleatoria los cuatro dígitos 1, 2, 3 y 4, ¿cuál es la probabilidad de que al menos un dígito ocupe su propio lugar?*
Se dice que el dígito i ocupa su propio lugar cuando éste fue el i-ésimo dígito escrito.
Como el espacio muestral Ω *consiste de todas las permutaciones de los dígitos* $\{1, 2, 3, 4\}$*, definamos para* $i = 1, 2, 3, 4$*, los cuatro siguientes eventos*

A_i = *{Subconjunto de* Ω *donde el dígito i ocupa su propio lugar}.*

Hallemos para $1 \le i < j < k \le 4$, $P(A_i)$, $P(A_i \cap A_j)$, $P(A_i \cap A_j \cap A_k)$.
Usando el Principio de la multiplicación obtengamos algunas de estas probabilidades

$i = 2$	*lugar*			
Etapas	i	1	3	4
N_0 *formas*	**1**	3	2	1
$\#A_2 = 3!$				

$i = 2$ $j = 4$	*lugar*			
Etapas	i	j	1	3
N_0 *formas*	**1**	**1**	2	1
$\#(A_2 \cap A_4) = 2!$				

$i=2\ j=3\ k=4$	*lugar*			
Etapas	i	j	k	1
N_0 *formas*	**1**	**1**	**1**	1
$\#(A_2 \cap A_3 \cap A_4) = 1!$				

Etapas	1	2	3	4
N_0 *formas*	1	1	1	1
$\#(A_1 \cap A_2 \cap A_3 \cap A_4) = 0!$				

Nos piden hallar $P(A_1 \cup A_2 \cup A_3 \cup A_4)$, *por Teorema 2.1.7,*

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = \sum_{i=1}^{4} P(A_i) - \sum_{1\le i<j\le 4} P(A_i \cap A_j)$$

$$+ \sum_{1\le i<j<k\le 4} P(A_i \cap A_j \cap A_k) - P(A_1 \cap A_2 \cap A_3 \cap A_4) = \sum_{i=1}^{4} \frac{\#A_i}{\#\Omega}$$

$$- \sum_{1\le i<j\le 4} \frac{\#(A_i \cap A_j)}{\#\Omega} + \sum_{1\le i<j<k\le 4} \frac{\#(A_i \cap A_j \cap A_k)}{\#\Omega} - \frac{\#(A_1 \cap A_2 \cap A_3 \cap A_4)}{\#\Omega}$$

$$= \sum_{i=1}^{4} \frac{3!}{4!} - \sum_{1\le i<j\le 4} \frac{2!}{4!} + \sum_{1\le i<j<k\le 4} \frac{1!}{4!} - \frac{0!}{4!} = \binom{4}{1}\frac{3!}{4!} - \binom{4}{2}\frac{2!}{4!} + \binom{4}{3}\frac{1!}{4!} - \binom{4}{4}\frac{0!}{4!}$$

$$= \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!}$$

Generalizando: Para toda $n \ge 2$: $P(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} (-1)^{(k+1)} \dfrac{1}{k!}$

Más aún: $\lim_{n\to\infty} P(\cup_{k=1}^{n} A_k)^c = e^{-1}$

2.1.4. Independencia y Probabilidad Condicional

Definición 2.1.12. *Sean* $(P, \Omega, \mathscr{F})$ *un espacio de probabilidad, y* $A, B \in \mathscr{F}$.

1. A, B *son eventos independientes si*

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

2. *La probabilidad de que ocurra el evento* A *dado que ocurrió* B, *denotada por* $P(A|B)$, *viene dada por*

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

La demostración del siguiente Teorema se deja como ejercicio al lector.

Teorema 2.1.13. *Sean* $(P, \Omega, \mathscr{F})$ *un espacio de probabilidad,*
$B \in \mathscr{F}$ *con* $P(B) > 0$, *entonces*

$$\begin{array}{rcl} P(\cdot|B) : \mathscr{F} & \longrightarrow & \mathbb{R} \\ A & \longmapsto & P(A|B) \end{array}$$

es una función de probabilidad.

Teorema 2.1.14. *Sean $(P, \Omega, \mathscr{F})$ un espacio de probabilidad, y $A, B \in \mathscr{F}$ con $0 < P(B) < 1$. Las afirmaciones siguientes son equivalentes*

1.	*A y B son independientes.*		
2.	$P(A\|B) = P(A)$.	2′.	$P(B\|A) = P(B)$.
3.	$P(A^c\|B) = P(A^c)$.	3′.	$P(B^c\|A) = P(B^c)$.
4.	$P(B\|A^c) = P(B)$.	4′.	$P(A\|B^c) = P(A)$.
5.	$P(B^c\|A^c) = P(B^c)$.	5′.	$P(A^c\|B^c) = P(A^c)$.

He aquí el significado de independencia entre eventos: Dos eventos son independientes si la ocurrencia o no, de uno de ellos, no altera la probabilidad de ocurrencia o no, del otro.

Demostración.

$$
\begin{aligned}
1. \Longleftrightarrow 2. \quad & P(A \cap B) = P(A)P(B) \Longleftrightarrow \frac{P(A \cap B)}{P(B)} = P(A) \Longleftrightarrow P(A|B) = P(A) \\
2. \Longleftrightarrow 3. \quad & P(A|B) = P(A) \Longleftrightarrow 1 - P(A|B) = 1 - P(A) \Longleftrightarrow P(A^c|B) = P(A^c) \\
3. \Longleftrightarrow 4. \quad & P(A^c|B) = P(A^c) \Longleftrightarrow \frac{P(A^c \cap B)}{P(B)} = P(A^c) \Longleftrightarrow \frac{P(B \cap A^c)}{P(A^c)} = P(B) \\
& \Longleftrightarrow P(B|A^c) = P(B) \\
4. \Longleftrightarrow 5. \quad & P(B|A^c) = P(B) \Longleftrightarrow 1 - P(B|A^c) = 1 - P(B) \\
& \Longleftrightarrow P(B^c|A^c) = P(B^c)
\end{aligned}
$$

Análogamente, intercambiando B con A en la demostración anterior se obtiene que $1., 2'.3'.4'.$ y $5'.$ son equivalentes.

□

Corolario 2.1.15. *Si una de las siguientes parejas de conjuntos son independientes, entonces las demás lo serán* $\{A, B\},\ \{A, B^c\},\ \{A^c, B\},\ \{A^c, B^c\}$.

Nota 2.1.16. *Los resultados de lanzamientos de monedas y dados son eventos independientes.*

Ejemplo 2.1.17. *Dos personas lanzan cada una, tres monedas equilibradas. ¿Cuál es la probabilidad de que ambas personas obtengan la misma cantidad de soles?*
*Como Ω consiste de las permutaciones de los 6 resultados (águila $\equiv$ **a** o sol $\equiv$ **s**) de las 2 personas, donde los 3 primeros pertenecen a la persona A, y los otros 3 a la persona*

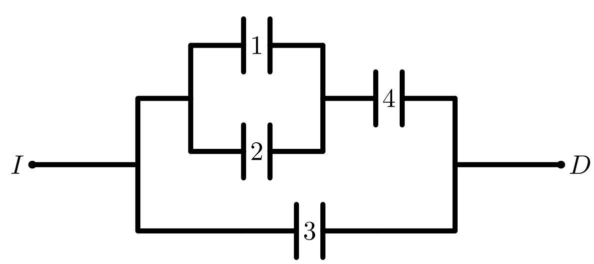
B, por el principio de múltiplicación: $\#\Omega = 2^6 = 64$.
Para $0 \leq i \leq 3$, sean A_i ={A obtuvo i resultados ***s***}, *B_i ={B obtuvo i resultados* ***s***}.
Nos piden hallar la probabilidad de la unión ajena $\cup_{i=1}^{3}(A_i \cap B_i)$.
Obsérvese que en el evento A_i sus 3 primeros elementos son permutaciones de una palabra con 3 letras de las cuales i de ellas son ***s*** *y las otras $3-i$ son* ***a***, *los últimos 3 elementos son letras arbitrarias de* ***s*** *o* ***a***, *por el principio de la múltiplicación $\#A_i = \dfrac{3!}{i!\,(3-i!)}\,2^3$, análogamente se obtiene $\#B_i = 2^3\,\dfrac{3!}{i!\,(3-i!)}$.*

$$P(\cup_{i=0}^{3}(A_i \cap B_i)) = \sum_{i=0}^{3} P(A_i \cap B_i) = \sum_{i=0}^{3} P(A_i)P(B_i) \qquad (\textit{Por independencia})$$

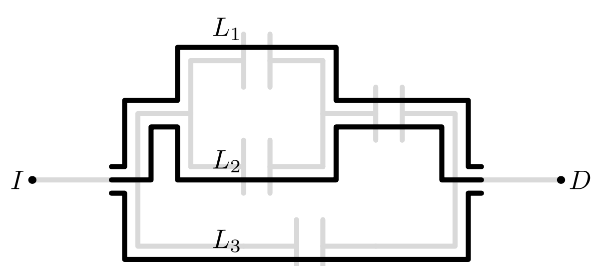
$$= \sum_{i=0}^{3} \left(\frac{\#A_i}{\#\Omega}\right)\left(\frac{\#B_i}{\#\Omega}\right)^2 = \sum_{i=0}^{3} \left(\frac{2^3}{2^6}\;\frac{3!}{i!\,(3-i!)}\right)^2 = \frac{\mathbf{1+9+9+1}}{\mathbf{64}} = \frac{5}{16}.$$

Obsérvese : $\#(A_i \cap B_i) = 20$ con probabilidad c/u de $1/64$

Ejemplo 2.1.18. *En la siguiente figura se supone que la probabilidad de que cada relevador esté cerrado es p y que cada relevador se abre o se cierra independiente de cualquier otro. Encontrar la probabilidad de que la corriente pase de I a D.*



Obsérvese que la corriente pasa de I a D por alguna de las tres lineas L_1, L_2 o L_3,



Para $i = 1, 2, 3$, sean C_i la corriente pasa de I a D por la linea L_i. Como la corriente pasa de I a D, si al menos pasa por una de estas tres lineas, entonces, nos piden

$P(C_1 \cup C_2 \cup C_3)$. *Además para* $i = 1, 2, \ldots, 6$, *sean* R_i *el relevador* i *está cerrado.*

$$\begin{aligned} &P(C_1 \cup C_2 \cup C_3) = P(C_1) + P(C_2) + P(C_3) - P(C_1 \cap C_2) - P(C_1 \cap C_3) + \\ &- P(C_2 \cap C_3) + P(C_1 \cap C_2 \cap C_3) = P(R_1 \cap R_4) + P(R_2 \cap R_4) + P(R_3) + \\ &- P(R_1 \cap R_2 \cap R_4) - P(R_1 \cap R_3 \cap R_4) - P(R_2 \cap R_3 \cap R_4) + P(R_1 \cap R_2 \cap R_3 \cap R_4) \\ &= p^2 + p^2 + p - p^3 - p^3 - p^3 + p^4 = p^4 - 3p^3 + 2p^2 + p \quad \textit{(por indep. de las } R_i\textit{)}. \end{aligned}$$

2.2. Teoremas Básicos

2.2.1. Teorema de la Multiplicación

Teorema 2.2.1. *Sean* $(P, \Omega, \mathscr{F})$ *un espacio de probabilidad, y* $A_1, A_2, \ldots, A_n \in \mathscr{F}$, *entonces*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)\ P(A_2|A_1) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Demostración. La demostración se hará por inducción matemática.
Paso 1. Probemos que el Teorema se cumple para un primer elemento, $n = 2$.
Por definición de Probabilidad Condicional, se sigue

$$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1).$$

Paso 2. Ahora supongamos que el Teorema se cumple para n=k

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)\ P(A_2|A_1) \cdots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}).$$

Paso 3. Finalmente demostremos que el Teorema se cumple para $n = k + 1$

$$P(A_1 \cap \cdots \cap A_{k+1}) = P((A_1 \cap A_2 \cap \cdots \cap A_k) \cap A_{k+1})$$

Por Paso 1 $\quad P(A_1 \cap \cdots \cap A_{k+1}) = P(A_1 \cap A_2 \cap \cdots \cap A_k)P(A_{k+1}|A_1 \cap A_2 \cap \cdots \cap A_k)$

Por la hipótesis de inducción que es el Paso 2, se obtiene lo que se quería demostrar

$$P(A_1 \cap \cdots \cap A_{k+1}) = (P(A_1) \cdots P(A_k|A_1 \cap \cdots \cap A_{k-1}))P(A_{k+1}|A_1 \cap \cdots \cap A_k).$$

□

Ejemplo 2.2.2. *De una urna que contiene 6 artículos defectuosos y 4 no defectuosos, se eligen 3 de ellos al azar sin sustitución, ¿cuál es la probabilidad de los 3 artículos seleccionados sean defectuosos?*
Ω *consiste de todas las permutaciones de 3 elementos tomados de la urna.*
Para $i = 1, 2, 3$, *sean* $D_i = \{w \in \Omega | \textit{el } i\textit{-ésimo artículo elegido es defectuoso}\}$
Nos piden hallar $P(D_1 \cap D_2 \cap D_3)$, *por el Teorema de la Multiplicación,*

$$P(D_1 \cap D_2 \cap D_3) = P(D_1)\ P(D_2|D_1)P(D_3|D_1 \cap D_2) = \left(\frac{6}{10}\right)\left(\frac{5}{9}\right)\left(\frac{4}{8}\right) = \frac{1}{6}$$

Definición 2.2.3. *Se dice que* $B_1, B_2, \ldots, B_n \in \mathscr{F}$ *es una partición de* Ω *si:*

1. $\cup_{i=1}^{n} B_i = \Omega$,
2. $B_i \cap B_j = \varnothing, \quad i \neq j$,
3. $P(B_i) > 0, \quad i = 1, 2, \ldots, n$.

2.2.2. Teorema de la Probabilidad Total

Teorema 2.2.4. *Sean* $(P, \Omega, \mathscr{F})$ *un espacio de probabilidad, y* $B_1, B_2, \ldots, B_n \in \mathscr{F}$ *partición de* Ω*, entonces para toda* $A \in \mathscr{F}$

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n).$$

Demostración.

$$P(A) = P(A \cap \Omega) = P(A \cap (\cup_{i=1}^{n} B_i)) = P(\cup_{i=1}^{n} (A \cap B_i))$$

como $(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \varnothing = \varnothing, \ i \neq j$, por axioma A_3,

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

por Teorema 2.2.1

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n).$$

□

Ejemplo 2.2.5. *Cuatro urnas llevan los números 1,2,3,4. La urna* i *contine* i *bolas blancas y* $4-i$ *bolas negras, con* $i = 1, 2, 3, 4$*. Se selecciona, al azar, una urna y después se saca una bola de dicha urna, ¿cuál es la probabilidad de que esta bola seleccionada sea negra?*
Sean, $U_i = \{w \in \Omega / \text{la bola es seleccionada de la urna } i\}$*, para* $i = 1, 2, 3, 4$*, y* $N = \{w \in \Omega / \text{la bola seleccionada es negra}\}$*, así para* $i = 1, 2, 3, 4$*,*

$$P(U_i) = \frac{1}{4}, \quad P(N|U_i) = \frac{4-i}{4}.$$

Como U_1, U_2, U_3, U_4 *es una partición de* Ω*, por teorema de la Probabilidad Total,*

$$\begin{aligned} P(N) &= P(N|U_1)P(U_1) + P(N|U_2)P(U_2) + P(N|U_3)P(U_3) + P(N|U_4)P(U_4) \\ &= \left(\frac{3}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{2}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{0}{4}\right)\left(\frac{1}{4}\right) = \frac{3}{8} \end{aligned}$$

2.2.3. Teorema de Bayes

Teorema 2.2.6. *Sean* $(P, \Omega, \mathscr{F})$ *un espacio de probabilidad, y* $B_1, B_2, \ldots, B_n \in \mathscr{F}$ *partición de* Ω*, entonces para toda* $A \in \mathscr{F}$

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)}.$$

Demostración.

Por defición, $P(B_j|A) = \dfrac{P(A \cap B_j)}{P(A)}$. Por Teoremas 2.2.1, 2.2.4,

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)}.$$

□

Ejemplo 2.2.7. *Un bolso contiene tres monedas, una de las cuales está acuñada con dos soles, mientras que las otras dos monedas son normales y no son irregulares. Se escoge una moneda al azar y se lanza cuatro veces en forma sucesiva. Si cada vez sale sol, ¿cuál es la probabilidad de que ésta sea la moneda con dos soles?*
Sean $N = \{w \in \Omega /$ *se utilizó la moneda normal*$\}$*, y* $A = \{s, s, s, s\}$.
Por independencia en los cuatro lanzamientos y usando Teorema de Bayes,

$$P(N^c|A) = \frac{P(A|N^c)\ P(N^c)}{P(A|N)\ P(N) + P(A|N^c)\ P(N^c)} = \frac{(1)^4\ (\frac{1}{3})}{(\frac{1}{2})^4\ (\frac{2}{3}) + (1)^4\ (\frac{1}{3})} = \frac{8}{9}.$$

Resumen de conceptos básicos
$P(\varnothing) = 0, \qquad 0 \leq P(A) \leq 1, \qquad P(\Omega) = 1, \qquad P(A^c) = 1 - P(A)$
<u>$P(A)$</u>
$P(A) = \dfrac{\#A}{\#\Omega}$ $\quad$ Ω finito equiprobable
$P(A) = P(A\mid B_1)P(B_1) + P(A\mid B_2)P(B_2) + \cdots + P(A\mid B_n)P(B_n)$ (Teorema de la probabilidad total)
<u>Diferencia: $A - B = A \cap B^c$</u>
$P(A - B) = P(A) - P(B)$ (Cuando $B \subset A$)
$P(A - B) = P(A) - P(A \cap B)$
<u>Unión</u>
$P(A \cup B \cup C) = P(A) + P(B) + P(C)$ (Eventos mutuamente excluyentes a pares)
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B) \cap C)$
<u>Intersección</u>
$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ (Eventos Independientes)
$P(A \cap B \cap C) = P(A) \cdot P(B\mid A) \cdot P(C\mid A \cap B)$ (Teorema de la multiplicación)
<u>Probabilidad condicional</u>
$P(A\mid B) = P(A)$ (Eventos Independientes)
$P(A\mid B) \doteq \dfrac{P(A \cap B)}{P(B)}$
$P(B_k\mid A) = \dfrac{(A\mid B_k)P(B_k)}{P(A\mid B_1)P(B_1) + P(A\mid B_2)P(B_2) + \cdots + P(A\mid B_n)P(B_n)}$ (Teorema de Bayes)

2.3. Ejercicios

1. Una instalación consta de dos calderas y un motor. Sea A el evento de que el motor está en buenas condiciones, mientras que los eventos B_k $(k = 1, 2)$ son los eventos de que la k-ésima caldera está en buenas condiciones. El evento C es que la instalación pueda funcionar. Si la instalación funciona cada vez que le motor y al menos una caldera funcione, expresar C y C^c en términos de A y de los eventos B_i.

2. Un mecanismo tiene dos tipos de repuestos, digamos I y II. Suponga que hay dos del tipo I y tres del tipo II. Definimos los eventos A_k, $k = 1, 2$ y $B_j =$, $j = 1, 2, 3$ como sigue, $A_k =${la k-ésima unidad tipo I está funcionando correctamente}, $B_k =${la j-ésima unidad tipo II está funcionando correctamente}. Finalmente sea el evento $C =$ { el mecanismo funciona}. Dado que el mecanismo funciona si al menos una unidad del tipo I y dos unidades del tipo II funcionan, expresar el evento C en términos de las A_k y las B_k.

3. Una gran tienda vende camisas deportivas en tres tallas: pequeña, mediana y grande; tres modelos: a cuadros, estampada y de franjas; y dos tipos de manga: corta y larga. Las siguiente tabla muestra las proporciones de camisas vendidas que caen en las diferentes combinaciones de categorías.

	Manga corta			Manga larga		
	Cuadro	Estampada	Franjas	Cuadro	Estampada	Franjas
Pequeña	0.04	0.02	0.05	0.03	0.02	0.03
Mediana	0.08	0.07	0.12	0.10	0.05	0.07
Grande	0.03	0.07	0.08	0.04	0.02	0.08

   a) ¿Cuál es la probabilidad de que la siguiente camisa vendida sea mediana, de manga larga y estampada?

   b) ¿Cuál es la probabilidad de que la siguiente camisa vendida sea mediana y estampada?

   c) ¿Cuál es la probabilidad de que la siguiente camisa vendida sea de manga corta?

   d) ¿Cuál es la probabilidad de que la talla de la siguiente camisa vendida sea mediana?

   e) ¿Cuál es la probabilidad de que el modelo de la siguiente camisa vendida sea estampado?

   f) Dado que la camisa que se acaba de vender era a cuadros y de manga corta, ¿cuál es la probabilidad de que su talla fuera mediana?

   g) Dado que la camisa que se acaba de vender era a cuadros, ¿cuál es la probabilidad de que su talla fuera mediana?

4. Suponga que A y B son eventos para los cuales $P(A) = x$, $P(B) = y$ y $P(A \cap B) = z$. Exprezar cada una de las probabilidades siguientes en términos de x, y y z.

   a) $P(A^c \cup B^c)$ *b*) $P(A^c \cap B)$ *c*) $P(A^c \cup B)$ *d*) $P(A^c \cap B^c)$.

5. Suponga que A, B y C son eventos tales que $P(A) = P(B) = P(C) = \frac{1}{4}$, $P(A \cap B) = P(C \cap B) = 0$ y $P(A \cap C) = \frac{1}{8}$. Calcular la probabilidad de que al menos uno de los eventos A, B o C ocurra.

6. Cierto tipo de motor eléctrico falla por obstrucción de los cojinetes, por combustión del embobinado o por el desgaste de las escobillas. Suponga que la probabilidad de la obstrucción es el doble de la combustión, la cual es cuatro veces más probable que la inutilización de las escobillas. ¿Cuál es la probabilidad de que la falla sea por cada uno de esos tres tres mecanismos?

7. En una habitación 10 personas tienen insignias numeradas del 1 al 10. Se eligen tres personas al azar y se les pide que abandonen la habitación simultáneamente y se anotan los números de las insignias.

   a) ¿Cuál es la probabilidad de que el número menor de las insignias sea 5?

   b) ¿Cuál es la probabilidad de que el número mayor de las insignias sea 5?

8. Supóngase que se escriben los dígitos 1, 2, 3 y 4 en un orden aleatorio. ¿Cuál es la probabilidad de que al menos un dígito ocupe su propio lugar?

9. Se extraen dos esferas de una urna que contiene diez esferas numeradas del 1 al 10, ¿cuál es la probabilidad de que la suma sea 10?

10. Un lote consta de 10 artículos sin defecto, 4 con pequeños defectos y 2 con defectos graves. Se elige un artículo al azar. Encontrar la probabilidad de que:

    a) no tenga defectos,

    b) no tenga defectos graves,

    c) que no tenga defecto o que tenga defecto grave.

11. Supóngase que de 10 objetos se eligen 5 al azar, con sustitución. ¿Cuál es la probabilidad de que ningún objeto sea elegido más de una vez?

12. Una caja contiene esferas numeradas $1, 2, \ldots, 20$. Se escogen dos esferas al azar. Encontrar la probabilidad de que los números sobre las esferas sean enteros consecutivos, si:

    a) las esferas se escogen sin sustitución,

    b) las esferas se escogen con sustitución.

13. ¿Cuántos subconjuntos que contengan al menos un elemento se pueden formar de un conjunto de 100 elementos?

14. Entre los números $1, 2, \ldots, 500$ se escoge uno al azar. ¿Cuál es la probabilidad de que el número escogido sea divisible entre 6 y entre 10?

15. De 6 números positivos y 8 números negativos se eligen 4 números al azar (sin sustitución) y se multiplican. ¿Cuál es la probabilidad de que el producto sea un número positivo?

16. Un lote contiene 7 artículos. Si se sabe que 3 artículos son defectuosos y se inspeccionan al azar y en forma sucesiva, ¿cuál es la probabilidad de que de que el quinto artículo inspeccionado seal el último defectuoso en el lote?

17. Cinco números se escogen al azar con sustitución entre los dígitos $\{0, 1, 2, \ldots 9\}$. ¿Cuál es la probabilidad de que dos no sean iguales?

18. La urna U_1 contiene x esferas blancas y y rojas. La urna U_2 contiene z esferas blancas y v rojas. Se escoge una esfera al azar de la urna U_1 y se pone en la urna U_2. Entonces se escoge una esfera al azar de la urna U_2. ¿Cuál es la probabilidad de que esta esfera sea blanca?

19. Una caja contiene 4 tubos malos y 6 buenos. Se sacan dos a la vez. Se prueba uno de ellos y se encuentra que es bueno. ¿Cuál es la probabilidad de que el otro también sea bueno?

20. En el problema anterior los tubos se verifican sacando uno al azar, se prueba y se repite hasta que se encuentran los cuatro tubos malos. ¿Cuál es la probabilidad de encontrar el cuarto tubo malo,

    a) en la quinta prueba?

    b) en la décima prueba?

21. Supónagase que A y B son dos eventos independientes asociados con un experimento. Si la probabilidad de que A o B ocurra es 0.6, mientras la probabilidad de que A ocurra es igual a 0.4, determinar la probabilidad de que B ocurra.

22. Veinte artículos, 12 de los cuales son defectuosos y 8 no defectuosos, se inspeccionan uno después del otro. Si estos artículos se escogen al azar, ¿cuál es la probabilidad de que:

    a) los dos primeros artículos inspeccionados sean defectuosos?

    b) los dos primeros artículos inspeccionados sean no defectuosos?

    c) entre los dos primeros artículos inspeccionados halla uno defectuoso y uno no defectuoso?

23. Un lote contiene 10 artículos. Si se sabe que 3 de ellos son defectuosos y se inspeccionan al azar y en forma sucesiva, ¿cuál es la probabilidad de que el quinto artículo inspeccionado sea el último defectuoso en el lote?

24. Cinco números se escogen al azar con sustitución entre los 10 dígitos. ¿Cuál es la probabilidad de que dos no sean iguales.

25. Supóngase que tenemos las urnas U_1, U_2 y U_3 cada una con tres cajones. U_1 tiene una moneda de oro en cada uno de sus cajones, U_2 una de oro en un cajón y una de plata en cada uno de los otros, y U_3 tiene una moneda de plata en cada uno de sus cajones Se escoge una urna al azar, y de ella se escoge un cajón al azar. La moneda que se encontró en este cajón es de oro. ¿Cuál es la probabilidad de la moneda provenga de U_2

26. En una fábrica de pernos, las máquinas A, B y C fabrican 25 %, 35 % y 40 % de la producción total, respectívamente. De lo que producen A, B y C el 5 %, 4 % y 2 % respectívamente son pernos defectuosos.

    a) Se escoge un perno al azar, ¿cuál es la probabilidad de que provenga de la máquina A?

    b) Se escoge un perno al azar ¿cuál es la probabilidad de que Sea defectuoso?

    c) Se escoge un perno al azar y resulta ser un defectuoso. ¿Cuál es la probabilidad de que el perno provenga de la máquina A?

27. Suponga que tenemos dos urnas, 1 y 2, cada una con dos cajones. La urna 1 tiene una moneda de oro en un cajón y una de plata en el otro, mientras que la urna 2 tiene una moneda de oro en cada uno de los cajones. Se escoge una urna al azar, y de ésta se escoge un cajón al azar. La moneda encontrada en este cajón es de oro. ¿Cuál es la probabilidad de que la moneda provenga de la urna 2?

28. Un bolso contiene tres monedas, una de las cuales está acuñada con dos soles, mientras que las otras dos monedas son normales y no son irregulares. Se escoge una moneda al azar y se lanza cuatro veces en forma sucesiva. Si cada vez sale sol, ¿cuál es la probabilidad de que ésta sea la moneda con dos soles?

29. En tres cajas se colocan canicas rojas, blancas y azules, distribuidas de la siguiente manera:

	Canicas		
Cajas	rojas	blancas	azules
A	5	3	2
B	1	8	1
C	3	1	6

Se selecciona una caja al azar y de ella se saca una canica al azar, si la canica seleccionada es roja. ¿Cuál es la probabilidad de que la caja usada haya sido la caja C?

30. Una fábrica tiene tres máquinas A, B y C produciendo la misma pieza para televisor. La máquina A produce 60 % de las piezas con un 95 % de ellas perfectas, la máquina B produce 30 % con 80 % de ellas perfectas y la máquina C produce 10 % con 65 % perfectas. Si se selecciona una pieza al azar,

    a) ¿cuál es la probabilidad de que ésta sea defectuosa?

    b) y es defectuosa, ¿cuál es la probabilidad de que haya dido producida en la máquina A?

31. Sean A y B dos eventos asociados con un experimento. Suponga que
    $P(A) = 0.4,\ P(A \cup B) = 0.7$ y $P(B) = p$.
    Hallar el valor de p con el que:

    i) A y B son mutuamente excluyentes,

    ii) A y B son independientes.

    Sugerencia: Desarrollar $P(A \cup B)$

32. Tres componenetes de un mecanismo, digamos C_1, C_2 y C_3 están colocados en serie (en línea recta). Suponga que esos mecanismos están agrupados en orden aleatorio. Sea R el evento $\{C_2$ está a la derecha de $C_1\}$, y S el evento $\{C_3$ está a la derecha de $C_1\}$. ¿Los eventos R y S son independientes? ¿Por qué?

33. Se lanza un dado y de manera independiente se escoge al azar una carta de una baraja normal. ¿Cuál es la probabilidad de que:

    a) el dado muestre un número par y la carta sea un as?

    b) el dado muestre un número par o la carta sea un as?

34. Un número binario está compuesto sólo de los dígitos 0 y 1. (Por ejemplo, 101011). Esos números tienen un papel muy importante en el funcionamiento de computadoras electrónicas. Suponga que un número binario está formado por n dígitos. Suponga que la probabilidad de que aparezca un dígito incorrecto es p y que los errores en dígitos diferentes son independiente uno de otro. ¿Cuál es la probabilidad de formar un número incorrecto?

35. Se lanza un dado 5 veces. ¿Cuál es la probabilidad de que ”5“ salga al menos una vez?

36. Dos personas lanzan tres monedas regulares cada una. ¿Cuál es la probabilidad de que obtengan el mismo número de caras?

37. Se lanzan dos dados y puesto que las caras muestran números diferentes, ¿cuál es la probabilidad de que una cara sea 4?

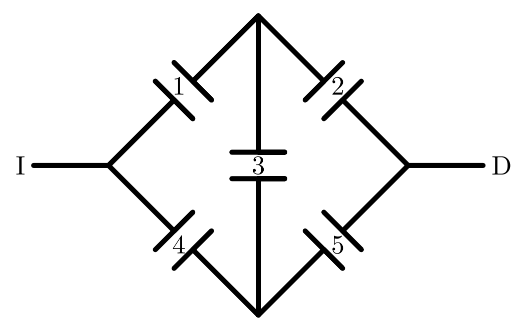
38. En la fabricación de cierto artículo se presenta un tipo de defectos con una probabilidad de 0.1, y un segundo tipo de defectos con probabilidad de 0.05. Suponiendo independencia entre estos tipos de defectos, ¿cuál es la probabilidad de que:

    a) un artículo no tenga ambas clases de defectos?

    b) un artículo sea defectuoso?

    c) suponiendo que un artículo sea defectuoso, tenga sólo un tipo de defecto?

39. En las figuras siguientes suponga que la probabilidad de que cada relevador esté cerrado es p y que cada relevador se abre o cierra independientemente de cualquier otro. Encontrar la probabilidad de que la corriente pase de I a D.



Capítulo 3

Variables aleatorias

Los datos estadísticos se pueden ver como valores de variables aleatorias, por lo que su análisis estadístico radica en el estudio de su variable aleatoria asociada. En este capítulo se estudia en general a las variables aleatorias (ver [2], [3]).

La σ-álgebra asociada (utilizada) para el conjunto de los números reales $\mathbb{R}$, es la mínima σ-álgebra que contiene a los intervalos de la forma $(-\infty, r)$, con $r \in \mathbb{R}$, la cual es conocida como σ-álgebra de Borel, y es denotada por $\mathscr{B}(\mathbb{R})$.

Definición 3.0.1. *Sea $(P, \Omega, \mathscr{F})$ un Espacio de Probabilidad.*
Una variable aleatoria (v. a.) X, es una función

$$\begin{array}{rcl} X : \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & X(\omega) \end{array}$$

tal que $X^{-1}(-\infty, r] = \{\omega \in \Omega \mid X(\omega) \leq r\} \in \mathscr{F}$, para toda $r \in \mathbb{R}$.

Denotaremos por $[X \leq x]$ al conjunto $\{\omega \in \Omega \mid X(\omega) \leq x\}$.

3.1. Función de distribución acumulada (fda)

Definición 3.1.1. *La Función de distribución acumulada o simplemente función de distribución F_X de la variable aleatoria X, se define como*

$$\begin{array}{rcl} F_X : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & F_X(x) = P[X \leq x]. \end{array}$$

Teorema 3.1.2. *[Continuidad secuencial monótana]*
Sea $\{A_n\} \in \mathscr{F}$ sucesión monótona, entonces:

1. *Si $A_n \subset A_{n+1}$ (denotaremos por$A_n \uparrow \cup_{i=1}^{\infty} A_i$), entonces $P(A_n) \uparrow P(\cup_{i=1}^{\infty} A_i)$.*

2. *Si $A_{n+1} \subset A_n$ (denotaremos por$A_n \downarrow \cap_{i=1}^{\infty} A_i$), entonces $P(A_n) \downarrow P(\cap_{i=1}^{\infty} A_i)$.*

Demostración. 1. Sean $B_1 = A_1,\ B_n = A_n - A_{n-1}, n \geq 2$ (**conjuntos disjuntos**), como $\cup_{i=1}^{n} B_i = A_n,\ \cup_{i=1}^{\infty} B_i = \cup_{i=1}^{\infty} A_i$, entonces

$$P(A_n) = P(\cup_{i=1}^{n} B_i) = \sum_{i=1}^{n} P(B_i) \quad \uparrow \quad \sum_{i=1}^{\infty} P(B_i) = P(\cup_{i=1}^{\infty} B_i) = P(\cup_{i=1}^{\infty} A_i).$$

2. Por Leyes de De Morgan $A_n^c \uparrow \cup_{n=1}^{\infty} A_n^c = (\cap_{n=1}^{\infty} A_n)^c$, por 1. se sigue

$$P(A_n) = 1 - P(A_n^c) \quad \downarrow \quad 1 - P((\cap_{i=1}^{\infty} A_i)^c) = P(\cap_{i=1}^{\infty} A_i).$$

□

Nota 3.1.3. *Por las propiedades de la función inversa: $f^{-1}(\cup_{i=1}^{\infty} A_i) = \cup_{i=1}^{\infty} f^{-1}(A_i)$, $f^{-1}(\cap_{i=1}^{\infty} A_i) = \cap_{i=1}^{\infty} f^{-1}(A_i)$, $f^{-1}(H - K) = f^{-1}(H) - f^{-1}(K)$, los eventos $[X < a]$ se comportan como intervalos (subconjuntos de $\mathbb{R}$).*

3.1.1. Propiedades de la Función de distribución acumulada

Teorema 3.1.4. *La función de distribución F_X de la variable aleatoria X satisface las siguientes propiedades:*

1. *Sean $a, b \in \mathbb{R}$ con $a < b$, entonces* $P[a < X \leq b] = F_X(b) - F_X(a)$.
2. *Es una función creciente.*
3. *Es continua por la derecha con límite por la izquierda, más aún, para $a \in \mathbb{R}$:*
   a) $\lim_{x \to a^+} F_X(x) = F_X(a)$.
   b) $\lim_{x \to a^-} F_X(x) = P[X < a]$.
4. $\lim_{x \to -\infty} F_X(x) = 0, \qquad \lim_{x \to \infty} F_X(x) = 1.$
5. *Para $a \in \mathbb{R}$* $P[X = a] = \lim_{x \to a^+} F_X(x) - \lim_{x \to a^-} F_X(x)$ ***(Salto de F_X en $x = a$)***
   $= F_X(a) - \lim_{x \to a^-} F_X(x)$.

Demostración.

Con $a < b$ se tiene que $[X \leq a] \subset [X \leq b]$, por Teorema 2.1.6:

1. $P[a < X \leq b] = P([X \leq b] - [X \leq a]) = P[X \leq b] - P[X \leq a]$
   $\therefore \quad P[a < X \leq b] = F_X(b) - F_X(a)$.
2. $P[X \leq a] \leq P([X \leq b] \qquad \therefore \quad F_X(a) \leq F_X(a)$.

3. *a*) Si $x_n \downarrow a$, entonces $[X \leq x_n] \downarrow \cap_{i=1}^{\infty}[X \leq x_i] = [X \leq a]$.

   Por Teorema3.1.2 $P[X \leq x_n] \downarrow P[X \leq a] \qquad \therefore \quad \lim_{x \to a^+} F_X(x) = F_X(a)$.

   b) Si $x_n \uparrow a$ entonces $[X \leq x_n] \uparrow \cup_{i=1}^{\infty}[X \leq x_i] = P[X < a]$.
   Por Teorema3.1.2 $P[X \leq x_n] \uparrow P[X < a] \qquad \therefore \quad \lim_{x \to a^-} F_X(x) = P[X < a]$.

4. Como en 3*a*), si $x_n \downarrow -\infty$, entonces $[X \leq x_n] \downarrow \varnothing$

   Por Teorema3.1.2 $P[X \leq x_n] \downarrow P(\varnothing) \qquad \therefore \quad \lim_{x \to -\infty} F_X(x) = 0$.

   Como en 3*b*), si $x_n \uparrow \infty$, entonces $[X \leq x_n] \uparrow \Omega$
   Por Teorema3.1.2 $P[X \leq x_n] \uparrow P(\Omega) \qquad \therefore \quad \lim_{x \to \infty} F_X(x) = 1$.

5. Si $x_n \uparrow a$ entonces $[x_n < X \leq a] \downarrow [X = a]$. Por 1. y Teorema 3.1.2

   $F_X(a) - P[X \leq x_n] = F_X(a) - F_X(x_n) = P[x_n < X \leq a] \downarrow P[X = a]$
   se sigue que $P[X \leq x_n] \uparrow F_X(a) - P[X = a]$

$$\begin{aligned} \therefore \quad \lim_{x \to a^-} F_X(x) &= F_X(a) - P[X = a] \\ &= \lim_{x \to a^+} F_X(x) - P[X = a] \quad \text{por 3.}a). \end{aligned}$$

□

De este Teorema deducimos la siguiente afirmación.

Corolario 3.1.5. *La función de distribución F_X de la variable aleatoria X es continua en $x = a$, si y solo si $P[X = a] = 0$.*

3.2. Variable aleatoria Discreta

Definición 3.2.1. *Diremos que la variable aleatoria X es discreta, si su función de distribución F_X toma un número finito o numerable de valores.*

Nota 3.2.2. *Sea $\{y_n \in [0,1] \mid y_n < y_{n+1},\ n \in \mathbb{Z}\}$ el conjunto de valores de la función de distribución discreta F_X, entonces por Teorema 3.1.4 existe $R_X = \{x_n \in \mathbb{R} \mid n \in \mathbb{Z}\}$, tales que:*

$F_X^{-1}(y_n) = [x_n, x_{n+1})$ *(por 2. y 3.)* F_X *es escalonada.*

$P[X = x_n] = y_n - y_{n-1}$ *(por 5.)*

$\lim_{n\to\infty} \sum_{k=-\infty}^{n} P[X = x_k] = \lim_{n\to\infty} y_n = 1$ *(por 4.)*
Por lo que la variable aleatoria X y su distribución F_X tienen la misma cantidad de valores. La gráfica de la distribución F_X es escalonada con conjunto de discontinuidad $R_X = \{x_n \in \mathbb{R} \mid n \in \mathbb{Z}\}$ y con longitud de salto $P[X = x_i]$.

3.2.1. Función de distribución de probabilidad (fdp)

Definición 3.2.3. *La función de distribución de probabilidad* f_X*, de la variable aleatoria* X *con valores en* $R_X = \{x_i \in \mathbb{R} \mid i \in \mathbb{Z}\}$*, se define como*

$$\begin{array}{rcl} f_X : R_X & \longrightarrow & \mathbb{R} \\ x_i & \longmapsto & f_X(x_i) = P[X = x_i]. \end{array}$$

Teorema 3.2.4. *La función de distribución de probabilidad* f_X *de la variable aleatoria discreta* X*, satisface las siguientes propiedades:*

1. $0 \leq f_X(x_i) \leq 1$
2. $\sum_{i=-\infty}^{\infty} f_X(x_i) = 1$

Nota 3.2.5. *Para una variable aleatoria discreta* X*, con valores* $R_X = \{x_i \mid i \in \mathbb{Z}\}$

1. $F_X(x) = \sum_{x_i \leq x} f_X(x_i), \qquad f_X(x_i) = F_X(x_i) - F_X(x_{i-1}).$
2. *La gráfica de su función de distribución* F_X *es una función escalonada.*

.

3.3. Variable aleatoria Continua

Definición 3.3.1.

1. *Diremos que la variable aleatoria* X *es continua, si su función de distribución* F_X *es continua.*
2. *Si además de ser continua la función* F_X*, es absolutamente continua (ver [6], [7]), entonces existe una función* f_X *llamada* ***función de densidad de probabilidad*** *o simplemente función de densidad de la variable aleatoria* X*, tal que*

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy.$$

Por Teorema Fundamental del Cálculo

$$f_X(x) = \frac{d[F_X(x)]}{dx}$$

3.3.1. Función de densidad de probabilidad (fdp)

Nota 3.3.2. *En este libro trabajaremos únicamente con variables aleatorias continuas, que tengan función de densidad de probabilidad.*

Teorema 3.3.3. *La función de densidad f_X de la variable aleatoria continua X, satisface las siguientes propiedades,*

1. $f_X(x) \geq 0$
2. $\int_{\mathbb{R}} f_X(x)\ dx = 1$

3.4. Variable aleatoria Mixta

Definición 3.4.1. *Diremos que la variable aleatoria X es mixta, si no es discreta ni continua, en este caso, la función de distribución F_X, tiene una parte discreta y otra parte continua por lo que su gráfica será discontinua no escalonada.*

Teorema 3.4.2. *La función de distribución F_X de una variable aleatoria mixta X, se puede escribir como combinación convexa de una función de distribución discreta $F^d_{X_1}$ de una variable aleatoria X_1, con otra función de distribución continua $F^c_{X_2}$ de una variable aleatoria X_2, es decir, existe $\alpha \in \mathbb{R}$, con $0 < \alpha < 1$, tal que*

$$F_X(x) = \alpha F^d_{X_1}(x) + (1-\alpha)F^c_{X_2}(x)$$

Demostración. Sea X_1 la variable aleatoria discreta con Rango $R_{X_1} = \{x_i \in \mathbb{R} \mid x_i$ es punto de discontinuidad de $F_X\}$, y función de distribución de probabildad

$$\begin{aligned} f^d_{X_1} : R_{X_1} &\longrightarrow \mathbb{R} \\ x_i &\longmapsto f_{X_1}(x_i) = \frac{1}{\alpha}P[X = x_i], \end{aligned}$$

donde $\alpha = \sum_{x_k \in R_X} P[X = x_k]$.
Por lo que su función de distribución es

$$F^d_{X_1}(x) = (1/\alpha) \sum_{x_k \leq x} P[X = x_k].$$

Sea X_2 la variable aleatoria continua con la siguiente función de distribución obtenida de F_X quitando sus saltos

$$F^c_{X_2}(x) = (1/(1-\alpha)) \left[\ F_X(x) - \sum_{x_k \leq x} P[X = x_k]\ \right] \quad (*)$$

multiplicando por $(1-\alpha)$ $\qquad (1-\alpha)F^c_{X_2}(x) = F_X(x) - \alpha(1/\alpha)\sum_{x_k \leq x} P[X = x_k]$

$$\therefore \quad F_X(x) = \alpha F^d_{X_1}(x) + (1-\alpha)F^c_{X_2}(x)$$

Derivando $(*)$ obtenemos la función de densidad de la variable aleatoria X_2

$$f^c_{X_2}(x) = \frac{dF^c_{X_2}(x)}{dx} = (1/(1-\alpha))\ \frac{dF_X(x)}{dx}$$

□

3.5. Esperanza de una variable aleatoria

Definición 3.5.1. *La Esperanza, Valor esperado, Media o Promedio, de la variable aleatoria X, denotada por μ_X o $E(X)$, se define como la siguiente Integral de Stieltjes (ver [6])*

$$\mu_X = E(X) = \int_{\mathbb{R}} x \; dF(x).$$

Más aún, si $g : \mathbb{R} \longrightarrow \mathbb{R}$ es función medible con la σ-álgebra de Borel $\mathscr{B}(\mathbb{R})$, es decir $g^{-1}(-\infty, r] \in \mathscr{B}(\mathbb{R})$ para toda $r \in \mathbb{R}$, entonces $g(X)$ es variable aleatoria con

$$E(g(X)) = \int_{\mathbb{R}} g(x) \; dF(x).$$

Así, en el caso de ser X variable aleatoria:

1. *Discreta,* $\quad \mu_X = \sum_{x_i \in R_X} x_i \; f_X(x_i), \qquad E(g(X)) = \sum_{x_i \in R_X} g(x_i) \; f_X(x_i)$

2. *Continua con función de densidad f_X*

$$\mu_X = \int_{\mathbb{R}} x \; f_X(x) \; dx \qquad E(g(X)) = \int_{\mathbb{R}} g(x) \; f_X(x) \; dx$$

3. *Mixta.* $\quad$ *Como $F_X(x) = \alpha F^d_{X_1}(x) + (1-\alpha) F^c_{X_2}(x)$, entonces*

$$\begin{aligned} E(g(X)) = \int_{\mathbb{R}} g(x) \; dF(x) =& \alpha \int_{\mathbb{R}} g(x_1) \; dF^d_{X_1}(x_1) + (1-\alpha) \int_{\mathbb{R}} g(x_2) \; dF^c_{X_2}(x_2) \\ =& \alpha E(g(X_1)) + (1-\alpha) E(g(X_2)) \end{aligned}$$

3.5.1. Propiedades del Valor esperado

Teorema 3.5.2. *Sean X, Y variables aleatorias. Si $a, b \in \mathbb{R}$, entonces:*

1. $E(a) = a$,

2. $E(aX + bY) = aE(X) + bE(Y)$.

3. *Si X, Y son independientes, entonces* $E(XY) = E(X)E(Y)$.

3.6. Varianza de una variable aleatoria

Definición 3.6.1. *La Varianza de la variable aleatoria X, denotada por σ_X^2 o $V(X)$, se define como*

$$\sigma_X^2 = V(X) = E(X - \mu_X)^2$$

Desarrollando el binomio al cuadrado y utilizando las propiedades del valor esperado

$$\sigma_X^2 = V(X) = E(X^2) - (\mu_X)^2$$

Así, en el caso de ser X variable aleatoria:

1. *Discreta* $\quad \sigma_X^2 = \sum_{x_i \in R_X} [(x_i)^2 \ f_X(x_i)] - (\mu_X)^2.$

2. *Continua con función de densidad f_X*

$$\sigma_X^2 = \int_{\mathbb{R}} x^2 \ f_X(x) \ dx - (\mu_X)^2.$$

3.6.1. Propiedades de la Varianza

Teorema 3.6.2. *Sean X, Y variables aleatorias. Si $a \in \mathbb{R}$, entonces:*

1. $V(a) = 0$,

2. $V(aX) = a^2 V(X)$,

3. $V(X + a) = V(X)$.

4. *Si X, Y son independientes, entonces* $V(X + Y) = V(X) + V(Y)$.

Ejemplo 3.6.3 (Caso Discreto)**.** *Considerando el experimento de lanzar dos dados*

$$\Omega = \{ \ (a, b) \mid 1 \leq a, b \leq 6 \ \}.$$

Sea X la siguiente variable aleatoria

$$\begin{array}{rccc} X: & \Omega & \longrightarrow & \mathbb{R} \\ & (a,b) & \longmapsto & a+b. \end{array}$$

cuyo Rango es $R_X = \{ \ 2, 3, \ldots, 11, 12 \ \}$.
Hallar: f_X, F_X, μ_X, σ_X^2, $P[4 < X < 8]$, $P[2.5 \leq X < 5.8]$

1. f_X viene dada por

$$f_X(x) = \begin{cases} \dfrac{k-1}{36}, & \text{si} \;\; 2 \leq k \leq 7, \\[2ex] \dfrac{13-k}{36}, & \text{si} \;\; 8 \leq k \leq 12. \end{cases}$$

k	2	3	4	5	6	7	8	9	10	11	12
$f_X(k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

A continuación obtenemos dos de sus valores

$$\begin{aligned} f_X(4) =& P[X=4] = P(\{(1,3);(2,2);(3,1)\}) = \frac{\#[X=4]}{\#\Omega} = \frac{3}{36} \\ f_X(9) =& P[X=9] = P(\{(3,6);(4,5);(5,4);(6,3)\}) = \frac{\#[X=9]}{\#\Omega} = \frac{4}{36} \end{aligned}$$

2. F_X viene dada por

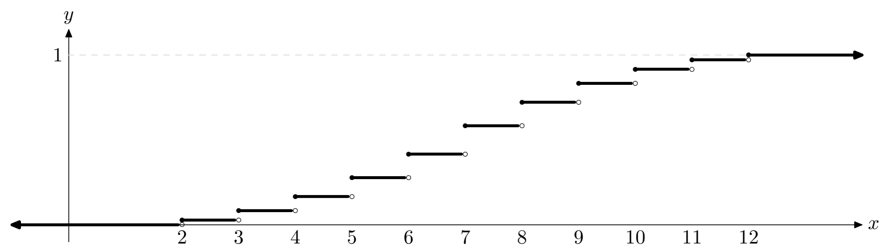
$$F_X(x) = \begin{cases} 0, & \text{si } x < 2, \\[2ex] \dfrac{1}{36}\dbinom{[\![x]\!]}{2}, & \text{si } 2 \leq x < 8, \\[2ex] 1 - \dfrac{1}{36}\dbinom{13-[\![x]\!]}{2}, & \text{si } 8 \leq x < 12, \\[2ex] 1, & \text{si } x \geq 12. \end{cases}$$

$[\![x]\!]$ es la parte entera del número real x, $\dbinom{a}{b}$ el coeficiente binomial.
ya que si $2 \leq x < 8$, por la propiedad de ser función acumulada de f_X

$$F_X(x) = \frac{1}{36}\sum_{k=1}^{[\![x]\!]-1} k = \frac{1}{36}\frac{([\![x]\!]-1)\;[\![x]\!]}{2} = \frac{1}{36}\binom{[\![x]\!]}{2}$$

Si $8 \leq x < 12$

$$F_X(x) = 1 - \frac{1}{36}\sum_{k=1}^{12-[\![x]\!]} k = 1 - \frac{1}{36}\frac{(12-[\![x]\!])\;13-[\![x]\!]}{2} = 1 - \frac{1}{36}\binom{13-[\![x]\!]}{2}$$



Gráfica de la función de distribución F_X.

Obtengamos la media μ_X y la varianza σ_X^2.

3. $\mu_X = \sum_{k=2}^{12} x_i \ f_X(x_i) = 2\left(\frac{1}{36}\right) + 3\left(\frac{2}{36}\right) + 4\left(\frac{3}{36}\right) + 5\left(\frac{4}{36}\right) + 6\left(\frac{5}{36}\right)$
$+ 7\left(\frac{6}{36}\right) + 8\left(\frac{5}{36}\right) + 9\left(\frac{4}{36}\right) + 10\left(\frac{3}{36}\right) + 11\left(\frac{2}{36}\right) + 12\left(\frac{1}{36}\right) = 7$

$E(X^2) = \sum_{k=2}^{12} (x_i)^2 \ f(x_i) = 4\left(\frac{1}{36}\right) + 9\left(\frac{2}{36}\right) + 16\left(\frac{3}{36}\right) + 25\left(\frac{4}{36}\right) + 36\left(\frac{5}{36}\right)$
$+ (49)\left(\frac{6}{36}\right) + 64\left(\frac{5}{36}\right) + 81\left(\frac{4}{36}\right) + 100\left(\frac{3}{36}\right) + 121\left(\frac{2}{36}\right) + 144\left(\frac{1}{36}\right) = \frac{329}{6}$

4. $\sigma_X^2 = E(X^2) - (\mu_X)^2 = \frac{329}{6} - 7^2 = \frac{35}{6}$

5. $P[4 < X < 8] = P[4 < X \leq 8] - P[X = 8] = (F_X(8) - F_X(4)) - f_X(8)$
$= (1/36)\left[\left(36 - \binom{13-8}{2}\right) - \binom{4}{2} - 5\right] = (1/36)[(36-10) - 6 - 5] = \frac{15}{36}$

6. $P[2.5 \leq X < 5.8] = P[2.5 < X \leq 5.8] + \cancel{P[X = 2.5]} - \cancel{P[X = 5.8]}$
$= F_X(5.8) - F_X(2.5) = (1/36)\left[\binom{5}{2} - \binom{2}{2}\right] = (1/36)[10 - 1] = \frac{9}{36} = \frac{1}{4}.$

Ejemplo 3.6.4 (Caso continuo). *La función de densidad f_X de la varible aleatoria X con $\mu_X = \frac{3}{5}$ está dada por*

$$f_X(x) = \begin{cases} a + bx^2, & si \quad 0 \leq x \leq 1, \\ 0 & en \ otro \ caso. \end{cases}$$

Hallar: a, b, F_X, y σ_X^2.

Resolvamos el siguiente sistema de dos ecuaciones con las incognitas a, b

$$\int_{\mathbb{R}} f_X(x)dx = 1 \qquad (1)$$
$$\int_{\mathbb{R}} x f_X(x)dx = \frac{3}{5} \qquad (2)$$

Simplifiquemos las ecuaciones (1) y (2)

$$1 = \int_{\mathbb{R}} f_X(x)dx = \int_0^1 (a + bx^2)dx = \left[ax + \frac{b}{3}x^3\right]_0^1 = a + \frac{b}{3} \tag{1}$$

$$\frac{3}{5} = \int_{\mathbb{R}} xf_X(x)dx = \int_0^1 x(a + bx^2)dx = \left[\frac{a}{2}x^2 + \frac{b}{4}x^4\right]_0^1 = \frac{a}{2} + \frac{b}{4} \tag{2}$$

Multiplicando por 3 la ecuación (1), y por 20 la ecuación (2), obtenemos el siguiente sistema de ecuaciones equivalente

$$3a + \ b = 3 \tag{3}$$

$$10a + 5b = 12 \tag{4}$$

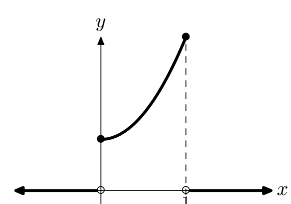
Multiplicando por 5 la ecuación (3) y restando la ecuación (4), obtenemos

$$5a = 3 \quad \Longrightarrow \quad a = \frac{3}{5}$$

Sustituyendo este valor de a en la ecuación (3), obtenemos

$$b = 3 - 3\left(\frac{3}{5}\right) \quad \Longrightarrow \quad b = \frac{6}{5}$$

$$\therefore \quad f_X(x) = \begin{cases} 0.6 + 1.2x^2, & si \quad 0 \leq x \leq 1, \\ 0, & en\ otro\ caso. \end{cases}$$

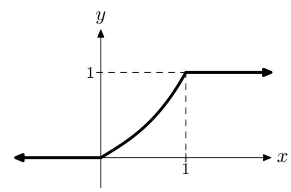


Gráfica de la función de densidad f_X.

Obtengamos la función de distribución F_X, por definición 3.3.1.

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\ dy = \begin{cases} 0, & si \quad x < 0, \\ \displaystyle\int_0^x (0.6 + 1.2y^2)\ dy, & si \quad 0 \leq x \leq 1 \\ 1, & si \quad x > 1. \end{cases}$$

$$\therefore \quad F_X(x) = \begin{cases} 0, & si \quad x < 0, \\ 0.6x + 0.4x^3, & si \quad 0 \leq x \leq 1 \\ 1, & si \quad x > 1. \end{cases}$$



Gráfica de la función de distribución F_X.

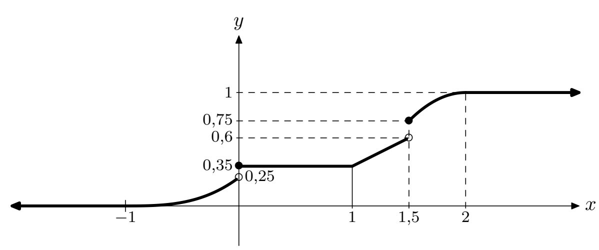
Finalmente obtegamos la varianza σ_X^2. Por Definición 3.6.1,2.

$$E(X^2) = \int_{\mathbb{R}} x^2\ f_X(x)\ dx = \int_0^1 x^2\ (0.6 + 1.2x^2)\ dx = \left[0.2x^3 + 0.24x^5\right]_0^1 = 0.44$$
$$\therefore \quad \sigma_X^2 = E(X^2) - (\mu_X)^2 = 0.44 - (0.6)^2 = 0.08$$

Ejemplo 3.6.5 (Caso mixto)**.** *Sea X variable aletoria con función de distribución F_X dada por*

$$F_X(x) = \begin{cases} 0, & si \quad x < -1, \\ 0.25(x+1)^3, & si \quad -1 \leq x < 0, \\ 0.35, & si \quad 0 \leq x < 1, \\ 0.5(x-1) + 0.35, & si\ 1 \leq x < 1.5, \\ -(x-2)^2 + 1, & si\ 1.5 \leq x < 2, \\ 1, & si\ x \geq 2. \end{cases}$$

Hallar: 1. $P[X = 0]$, 2. $P[X = 1.8]$, 3. $P[0 \leq X < 1.8]$, 4. $P[X \geq 0]$, 5. μ_X, 6. σ_X^2.



Gráfica de la distribución F_X, de la variable aleatoria mixta X.

En general mi estrategia para hallar probabildades, es escribir todo intervalo en función de conjuntos de la forma $[a < X \leq b]$ y $[X = c]$, para luego utilizar las tres siguientes afirmaciones:
- Por Teorema 2.1.6. Si $A \subset B$, entoces $P(B - A) = P(B) - P(A)$.
- Por Teorema 3.1.4: Longitud del salto en la discontinidad de F_X en $x = a$
$$P[X = a] = F_X(a) - \lim_{x \to a^-} F_X(x).$$
- Para toda $a < b$, se cumple $P[a < X \leq b] = F_X(b) - F_X(a)$.

1. $P[X = 0] = F_X(0) - \lim_{x \to 0^-} F_X(x) = 0.35 - 0.25 = 0.1$,
2. $P[X = 1.8] = F_X(1.8) - \lim_{x \to 1.8^-} F_X(x) = 0 \quad$ (pto. de cont. de F_X),
3. $P[0 < X \leq 1.8] = F_X(1.8) - F_X(0) = 0.96 - 0.35 = 0.61$,

Escribiendo $[0 \leq X < 1.8] = ([0 < X \leq 1.8] - [X = 1.8]) \cup [X = 0]$, por 1, 2, y 3.

$$P[0 \leq X < 1.8] = P[0 < X \leq 1.8] - P[X = 1.8] + P[X = 0] = .61 - 0 + .1 = 0.71$$

4. $P[X \geq 0] = 1 - P[X < 0] = 1 - P\big([X \leq 0] - [X = 0]\big)$
$$= 1 - \big(P([X \leq 0] - P[X = 0]\big) = 1 - F_X(0) + P[X = 0] = 0.75$$

Para hallar μ_X, y σ^2_X, obtengamos $f^d_{X_1}$ de la parte discreta, y $F^c_{X_2}$ de la parte continua de la variable aleatoria mixta X. Por Teorema 3.4.2

$$R_{X_1} = \{0, 1.5\}, \quad \alpha = P[X = 0] + P[X = 1.5] = 0.1 + 0.15 = 0.25 = 1/4$$

$$f^d_{X_1} : R_{X_1} \longrightarrow \mathbb{R}$$

$x_i \in R_{X_1}$	0	1.5
$f^d_{X_1}(x_i) = \frac{1}{\alpha} P[X = x_i]$	$4(0.1) = 0.4$	$4(0.15) = 0.6$

$$f^c_{X_2}(x) = (1/(1-\alpha))\ \frac{dF_X(x)}{dx} = \begin{cases} 0, & \text{si} \quad x < -1, \\ (x+1)^2, & \text{si} \quad -1 \leq x < 0, \\ 0, & \text{si} \quad 0 \leq x < 1, \\ \frac{2}{3}, & \text{si } 1 \leq x < 1.5, \\ -\frac{8}{3}(x-2), & \text{si } 1.5 \leq x < 2, \\ 0, & \text{si } x \geq 2. \end{cases}$$

$$
\begin{aligned}
E(X_1) &= \sum_{x_i \in R_{X_1}} x_i \ f^d_{X_1}(x_i) = 0 \ P[X_1 = 0] + 1.5 \ P[X_1 = 1.5] = (1.5)(0.6) = 0.9 \\
E[(X_1)^2] &= \sum_{x_i \in R_{X_1}} (x_i)^2 \ f^d_{X_1}(x_i) = 0^2 \ P[X_1 = 0] + (1.5)^2 \ P[X_1 = 1.5] = 1.35 \\
E(X_2) &= \int_{\mathbb{R}} x \ f^c_{X_2}(x) \ dx = \int_{-1}^{0} x(x+1)^2 \ dx + \tfrac{2}{3}\int_{1}^{1.5} x \ dx - \tfrac{8}{3}\int_{1.5}^{2} x(x-2) \ dx \\
&= \left[\frac{x^4}{4} + \frac{2}{3}x^3 + \frac{x^2}{2}\right]_{-1}^{0} + \frac{1}{3}\left[x^2\right]_1^{1.5} - \frac{8}{3}\left[\frac{x^3}{3} - x^2\right]_{1.5}^{2} = -\frac{1}{12} + \frac{5}{12} + \frac{5}{9} = \frac{8}{9}. \\
E[(X_2)^2] &= \int_{\mathbb{R}} x^2 \ f^c_{X_2}(x) \ dx = \int_{-1}^{0} x^2(x+1)^2 \ dx + \tfrac{2}{3}\int_{1}^{1.5} x^2 \ dx - \tfrac{8}{3}\int_{1.5}^{2} x^2(x-2) \ dx \\
&= \left[\frac{x^5}{5} + \frac{2}{4}x^4 + \frac{x^3}{3}\right]_{-1}^{0} + \frac{2}{9}\left[x^3\right]_1^{1.5} - \frac{8}{3}\left[\frac{x^4}{4} - \frac{2}{3}x^3\right]_{1.5}^{2} = \frac{1}{30} + \frac{19}{36} + \frac{67}{72} \\
&= \frac{179}{120}.
\end{aligned}
$$

Obtengamos la media μ_X, y la esperanza $E[(X)^2]$. Por Definición 3.5.1,3.

5. $\mu_X = E(X) = \alpha E(X_1) + (1-\alpha)E(X_2) = 0.25(0.9) + 0.75\left(\frac{8}{9}\right) = \frac{107}{120}.$

$$E[(X)^2] = \alpha E[(X_1)^2] + (1-\alpha)E[(X_2)^2] = 0.25(1.35) + 0.75\left(\frac{179}{120}\right) = \frac{233}{160}.$$

Finalmente obtengamos la varianza σ^2_X. Por Definición 3.6.1,2.

6. $\sigma^2_X = E[(X)^2] - (\mu_X)^2 = \frac{233}{160} - \left(\frac{107}{120}\right)^2 = 0.66118$

3.7. Ejercicios

1. Dado que la V.A. discreta X tiene la fda

$$F_X(x) = \begin{cases} 0, & \text{para } x < -1, \\ 1/4, & \text{para } -1 \le x < 1, \\ 1/2, & \text{para } 1 \le x < 3, \\ 3/4, & \text{para } 3 \le x < 5, \\ 1, & \text{para } 5 \le x. \end{cases}$$

Obtener la gráfica de F y

a. $P[X < 3]$.
b. $P[X = 3]$.
c. $P[X \leq 3]$.
d. $P[-0.4 < X < 4]$.
e. $P[X \geq 1]$.
f. $P[X = 5]$.

2. Sea X una variable aleatoria continua con fdp f dada por

$$f(x) = \begin{cases} ax, & 0 \leq x \leq 1, \\ a, & 1 < x \leq 2, \\ -ax + 3a, & 2 < x \leq 3, \\ 0, & \text{en otro caso.} \end{cases}$$

   a. Determinar la constante a.
   b. Determinar F y dibujar su gráfica.
   c. Hallar $P[X \geq 2]$.

3. Si la fdp de la variable aleatoria continua Y está dada por

$$f(y) = \begin{cases} \dfrac{c}{\sqrt{y}}, & 0 \leq y \leq 4, \\ 0, & \text{en otro caso.} \end{cases}$$

   Obtener

   a. El valor de c.
   b. La fda F.
   c. $P[Y > 1]$.

4. Sea f la función definida por,

$$f(x) = \begin{cases} x + 1, & \text{si } -1 < x < 0, \\ 2x - 6, & \text{si } 3 < x < c, \\ 0, & \text{en otro caso.} \end{cases}$$

   donde c es una constante.

   a. Determinar el valor de c para que f, sea una fdp.
   b. $P[0 < X < 3]$.
   c. $P[X \leq 3]$.

Capítulo 4

Variables aleatorias discretas de uso común

En este capítulo estudiamos a las variables aleatorias discretas que más se usan.
En todas ellas salvo en la última se determina el número de éxitos en cierta cantidad de observaciones (observaciones donde hay sólo dos posibles resultados: éxito y fracaso), por ejemplo en una línea de producción, la cantidad de artículos defectuosos en los primeros 100 artículos producidos.
La última es la variable aleatoria Poisson que determina el número de éxitos ocurridos en un intervalo, por ejemplo la cantidad de clientes que arriban a un banco en cierto intervalo de tiempo (ver [2], [3]).

Definición 4.0.1. *[independencia] Diremos que las variables aleatorias discretas X, Y son independientes si,*

$$P([X \in A] \cap [Y \in B]) = P[X \in A]\ P[Y \in B], \qquad \textit{(ver definición 2.1.12)},$$

donde $A \subset R_X,\ B \subset R_Y$.
En particular, si $x_i \in R_X,\ y_i \in R_Y$: $\quad P([X = x_i], [Y = y_j]) = f_X(x_i) f_Y(y_i)$.

4.1. Distribución Bernoulli

Definición 4.1.1 (Ensayo de Bernoulli)**.** *Un ensayo de Bernoulli con parámetro p es el experimento aleatorio más simple, el cual consiste de dos posibles resultados a los que identificaremos como éxito E, y fracaso F, y donde $P(E) = p$.*

Definición 4.1.2 (Distribución de Bernoulli)**.** *La variable aleatoria asociada al ensayo de Bernoulli es la biyección,*

$$\begin{aligned} X : \Omega &\longleftrightarrow \{0,1\} \\ \omega &\longmapsto X(\omega) = \begin{cases} 0, & si\ \ \omega = F \\ 1, & si\ \ \omega = E. \end{cases} \end{aligned}$$

la cual es conocida como, variable aleatoria con distribución Bernoulli con parámetro p, esto se escribe como, $X \sim B_{er}(p)$, donde $p = P(E) = P[X = 1]$, luego su función de distribución de probabilidad f_X viene dada por,

$$f_X(x_i) = \begin{cases} p, & si \;\; x_i = 1 \\ q, & si \;\; x_i = 0. \end{cases}$$

donde $q = 1 - p$

Teorema 4.1.3. *Sea X una variable aletoria con distribución Bernoulli con parámetro p; entonces,*

1. $E(X) = p$
2. $V(X) = pq$

Demostración.

$$\begin{aligned} &E(X) = 0(q) + 1(p) = p. \\ &E[(X)^2] = 0^2(q) + 1^2(p) = p. \\ &V(X) = E[(X)^2] - (E(X))^2 = p - p^2 = p(1-p) = pq. \end{aligned}$$

□

4.2. Distribución Binomial

Definición 4.2.1 (Distribución Binomial)**.** *La variable aleatoria discreta X, que determina el número de éxitos en un experimento que consiste de n ensayos independientes de Bernoulli con el mismo parámetro $p = P(E_i)$, para $i = 1, 2, \ldots, n$, es conocida como variable aleatoria con distribución Binomial con parámetros n, p, lo cual escribiremos como $X \sim B(n, p)$.*

Nota 4.2.2. *Como para $(1 \leq i \leq n)$, el i-ésimo ensayo independiente de Bernoulli con parámetro p, tiene asociada la varible aleatoria independiente de Bernoulli X_i con parámetro p, entonces toda variable aleatoria con distibución Binomial con parámetros n, p, se puede escribir como suma de estas n variables aleatorias independientes con distribución Bernoulli con el mismo parámetro p, es decir,*

$$X = X_1 + \cdots + X_n, \quad X_i \sim B_{er}(p), \;\; i = 1, \ldots, n \;\; independientes.$$

Teorema 4.2.3. *La función de distribución de probabilidad f_X, de la variable alatoria X con distribución Binomial con parámetros n, p, viene dada por:*

$$f_X(k) = \binom{n}{k} \; p^k \; q^{n-k}, \quad para \;\; k \in R_X = \{0, 1, \ldots n\}.$$

Demostración.

Sea $k \in R_X$, entonces,

$$\begin{aligned} f_X(k) &= P[X = k] \\ &= P\{\text{permutaciones de la palabra } (\underbrace{E \cdots E}_{k}\underbrace{F \cdots F}_{n-k})\} \end{aligned}$$

Por independencia, por conmutativadad del producto, y por la fórmula de permutación de palabras Teorema 1.4.1,

$$f_X(k) = \frac{n!}{n_E! \cdot n_F!}(P(E))^k(P(F))^{n-k} = \frac{n!}{k! \cdot (n-k)!}\ p^k(1-p)^{n-k} = \binom{n}{k}\ p^k\ q^{n-k}$$

$\square$

Teorema 4.2.4. *Sea X una variable aletoria con distribución Binomial con parámetros $n,\ p$; entonces,*

1. $E(X) = np$

2. $V(X) = npq$

Demostración.

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k\ f(k) = \sum_{k=0}^{n} k\ \frac{n!}{k!\ (n-k)!}\ p^k\ q^{n-k} \\ &= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\ ((n-1)-(k-1))!}\ p^{k-1}\ q^{(n-1)-(k-1)} \end{aligned}$$

Por el cambio de variable $j = k - 1$, y por Teorema 3.2.4,2; con f_Y fdp. de la v.a. $Y \sim B(n-1, p)$.

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j}\ p^j\ q^{(n-1)-j} = np \sum_{y_i \in R_Y} f_Y(y_i) = np.$$

Análogamente, considerando $j = k-2$ y una v.a. $Y \sim B(n-2, p)$,

$$\begin{aligned}
E[X(X-1)] &= \sum_{k=0}^{n} k(k-1) \ \frac{n!}{k! \ (n-k)!} \ p^k \ q^{n-k} \\
&= n(n-1)p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)! \ ((n-2)-(k-2))!} \ p^{k-2} \ q^{(n-2)-(k-2)} \\
&= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} \ p^j \ q^{(n-2)-j} = n(n-1)p^2.
\end{aligned}$$

$E[X^2] = E[X(X-1)] + E(X) = n(n-1)p^2 + np = npq + n^2p^2.$
$V(X) = E(X^2) - (E(X))^2 = npq.$

Estos resultados se pueden obtener directamente por Nota 4.2.2 y Teorema 4.1.3.

$$\begin{aligned}
E(X) &= E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = p + \cdots + p = np. \\
V(X) &= V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n) = pq + \cdots + pq = npq.
\end{aligned}$$

□

Ejemplo 4.2.5. *La variable aleatoria que determina el número de esferas blancas, al extraer con reemplazo de manera aleatoria n esferas, de una urna que contiene N esferas en total, donde r de ellas son blancas, y el resto negras, tiene distribución Binomial con parámetros n, $p = r/N$.*

Tenemos n ensayos de Bernoulli, los cuales consisten en n extracciones con reemplazo de una esfera en esta urna. Cada uno con parámetro $p = P(E) = \frac{r}{N}$, donde éxito E consiste en extraer una esfera blanca de la urna.
Los n ensayos son independientes por Teorema 2.1.14, ya que, por el reemplazo,

$$P(\omega_i | \omega_j) = P(\omega_i)$$

donde para $1 \leq k \leq n$,

ω_k es el resultado (éxito o fracaso) de la k-ésima extracción con reemplazo.

Sea $X = \#$ de esferas blancas obtenidas en las n extracciones, entonces,
$X \sim B(n, \frac{r}{N})$, por lo que la probabilidad de que k de ellas sean blancas es,

$$f_X(k) = P[X = k] = \binom{n}{k} \left(\frac{r}{N}\right)^k \left(\frac{N-r}{N}\right)^{n-k}$$

Ejemplo 4.2.6. *Suponga que en cierta población, la probabilidad de que en un parto nazca un niño es 0.4, hallar la probabilidad de que en cinco partos en esta ciudad:*

1. *nazcan dos niños.*
2. *al menos nazcan una niña y un niño.*
3. *¿Cuántas niñas se espera que nazcan en estos cinco partos?*
4. *De 2 000 familias de esta población con tres hijos, ¿cuántas se espera que no tengan niños?*

Sea $X = \#$ de niñas en una familia de cinco hijos, entonces,

$$X \sim B(5, 0.6), \quad \text{con Éxito=nazca niña.}$$

Nos piden hallar: 1. $P[X = 3]$, 2. $P[1 \le X \le 4]$, 3. $E(X)$.

$$\begin{aligned}
&1. \quad P[X=3] = \binom{5}{3}(0.6)^3(0.4)^2 = 0.3456 \\
&2. \quad P[1 \le X \le 4] = 1 - P[X \in \{0,5\}] = 1 - P[X=0] - P[X=5] \\
&\qquad\qquad = 1 - \binom{5}{0}(0.4)^5 - \binom{5}{5}(0.6)^5 = 1 - 0.088 = 0.912 \\
&3. \quad E(X) = np = 5(0.6) = 3
\end{aligned}$$

4. Sea $Y = \#$ de familias con 3 niñas entre estas 2 000 familias con tres hijos

$Y \sim B(2\ 000, 0.216)$, con Éxito=familia con tres hijos todas niñas,
por lo que $p = P(E) = P(\text{En la familia halla tres niñas}) = (0.6)^3 = 0.216$.

$E(Y) = \#$ Esperado de familias con 3 niñas entre las 2 000 familias con 3 hijos

$$= 2\ 000(0.216) = 432$$

Ejemplo 4.2.7.

Para acreditar un curso, un estudiante responde un examen de opción múltiple con preguntas con cuatro opciones como respuesta, donde sólo una es la correcta. Suponga que con probabilidad 0.6 el estudiante conoce la respuesta correcta de cada una de las preguntas y que el estudiante selecciona la respuesta correcta cuando la conoce y en caso contrario, selecciona al azar una de las cuatro opciones. Si el examen consta de diez preguntas y si el curso se acredita con calificación mínima de siete, ¿cuál es la probabilidad de que este alumno logre acreditarlo y cuántas preguntas se espera que responda correctamente?

Sea $X = \#$ de preguntas del examen respondidas correctamente, entonces,

$$X \sim B(10, p), \quad \text{con Éxito=una de las preguntas es respondida correctamente,}$$

Sea B el Evento donde el alumno conoce la respuesta, por Teorema de la prob. total,

$$p = P(E) = P(E|B)P(B) + P(E|B^c)P(B^c) = 1(0.6) + (0.25)(0.4) = 0.7$$

La calificación mínima de siete se garantiza respondiendo correctamente siete o más preguntas. Nos piden hallar, $P[X \geq 7]$ y $E(X)$,

$$\begin{aligned} P[X \geq 7] &= \binom{10}{7}(0.7)^7(0.3)^3 + \binom{10}{8}(0.7)^8(0.3)^2 + \binom{10}{9}(0.7)^9(0.3) + \binom{10}{10}(0.7)^{10} \\ &= 0.6496 \\ E(X) &= 10(0.7) = 7 \end{aligned}$$

Por lo que se espera que se contesten correctamente siete preguntas.

4.3. Distribución Hipergeométrica

Definición 4.3.1. *La variable aleatoria que determina, el número de éxitos obtenidos al extraer sin reemplazo de manera aleatoria n esferas, de una urna que contiene N esferas en total, donde r de ellas son éxitos E, y el resto $N - r$ fracasos F, es conocida como, la variable aleatoria $X \sim H(N, n, r)$ con distribución Hipergeométrica, con parámetros n, r, N.*

Teorema 4.3.2 (Distribución Hipergeométrica)**.** *La variable aleatoria $X \sim H(r, n, N)$ con distribución Hipergeométrica, con parámetros n, r, N, tiene como función de densidad de probabilidad*

$$f_X(k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}, \quad \textit{para} \quad k = 0, 1, 2, \ldots, \text{mín}\{r, n\}.$$

Teorema 4.3.3. *La Esperanza y la Varianza de la variable aleatoria X, con distribución Hipergeométrica con parámetros r, n, N, vienen dadas por,*

1. $E(X) = np$

2. $V(X) = npq\ \dfrac{N-n}{N-1},$

donde: $p = \dfrac{r}{N}$, y $q = 1 - p.$

Demostración.

$$E[X^m] = \sum_{k=0}^{n} k^m\ f_X(k) = \sum_{k=1}^{n} k^m\ \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$$

Usando la identidad: $\quad k\binom{i}{k} = i\binom{i-1}{k-1}$

$$E[X^m] = \frac{nr}{N}\sum_{k=1}^{n} k^{m-1}\ \frac{\binom{r-1}{k-1}\binom{N-r}{n-k}}{\binom{N-1}{n-1}}$$

haciendo el cambio de variable, $k = j + 1$

$$\begin{aligned}
&= \frac{nr}{N}\sum_{j=0}^{n-1}(j+1)^{m-1}\ \frac{\binom{r-1}{j}\binom{N-r}{n-1-j}}{\binom{N-1}{n-1}} \\
&= \frac{nr}{N}E[(Y+1)^{m-1}]
\end{aligned}$$

donde $Y \sim H(r-1, n-1, N-1)$.

Con valores $m = 1,\ m = 2$, obtenemos,

$$\begin{aligned}
E(X) &= \frac{nr}{N}E[(Y+1)^0] = \frac{nr}{N} \\
E[X^2] &= \frac{nr}{N}E[(Y+1)] = \frac{nr}{N}\left[\frac{(n-1)(r-1)}{(N-1)} + 1\right] \\
V(X) &= E[X^2] - (E[X])^2 = \frac{nr}{N}\left[\frac{(n-1)(r-1)}{(N-1)} + 1 - \frac{nr}{N}\right] \\
&= \frac{nr}{N}\left[\frac{\cancel{nrN} - nN - rN + \cancel{N} + N^2 - \cancel{N} - \cancel{nrN} + nr}{N(N-1)}\right] = \frac{nr}{N}\frac{(N-r)(N-n)}{N(N-1)}
\end{aligned}$$

haciendo $p = \dfrac{r}{N}$, y $q = 1 - p$, obtenemos,

$$E(X) = np, \qquad V(X) = npq\,\frac{N-n}{N-1}$$

$\square$

Ejemplo 4.3.4. *En un estado de la República Mexicana se efectúa una lotería en la que se escogen al azar y sin reemplazo seis números entre 40. El boleto ganador será aquel que contenga estos seis números, y habrá reembolso si tiene al menos cuatro aciertos, es deicir si coinciden al menos con cuatro de estos números. Un jugador selecciona seis números antes de que el estado tome la muestra. ¿Cuál es la probabilidad de que el boleto del jugador:*

1. *sea el boleto ganador?*
2. *tenga reembolso?*
3. *¿Cuál es el número esperado de aciertos para este jugador?*

Sea X = número de aciertos del jugador = cantidad de números en el boleto del jugador, que coinciden con los 6 números seleccionados por el Estado, al azar y sin reemplazo entre 40.

$X \sim H(6, 6, 40)$, con Éxito=Acierto.

Nos piden hallar: 1. $P[X = 6]$, 2. $P[4 \leq X \leq 5]$, 3. $E(X)$.

1. $P[X = 6] = \dfrac{\binom{6}{6}\binom{34}{0}}{\binom{40}{6}} = \dfrac{1}{\binom{40}{6}} = 2.60526 \times 10^{-7}$

2. $P[4 \leq X \leq 5] = \dfrac{\binom{6}{4}\binom{34}{2}}{\binom{40}{6}} + \dfrac{\binom{6}{5}\binom{34}{1}}{\binom{40}{6}} = 0.002245$

3. $E(X) = 6\left(\dfrac{6}{40}\right) = 0.9$

Las variables aleatorias Binomial e Hipergeométrica son muy parecidas, ya que en ambas, se determina el número de éxitos obtenidos en n observaciones donde en cada una, se tiene los dos posibles resultados: Éxito y Fracaso. La única diferencia es que en la Binomial las n observaciones son independientes y en la Hipergeométrica las n observaciones (extracciones sin reemplazo) no son independientes, y como se muestra en el ejemplo 4.2.5, si estas n extracciones fueran con reemplazo se tendría independencia en las observaciones por lo que estaríamos en el caso de una distribución Binomial.

4.4. Distribución Geométrica

Definición 4.4.1. *La variable aleatoria X que determina, el número de ensayos sucesivos de Bernoulli, independientes y con el mismo parámetro p, hasta que ocurre el primer éxito E, es conocida como, la variable aleatoria $X \sim G(p)$ con distribución Geométrica, con parámetro p.*

Teorema 4.4.2 (Distribución Geométrica). *La variable aleatoria $X \sim G(p)$ con distribución Geométrica, con parámetro p, tiene como función de densidad de probabilidad*

$$f_X(k) = q^{k-1}p, \quad \textit{para} \ \ k = 1, 2, \ldots, n.$$

Teorema 4.4.3. *La Esperanza y la Varianza de la variable aleatoria X, con distribución Geométrica con parámetro p, vienen dadas por,*

1. $E(X) = \dfrac{1}{p}$
2. $V(X) = \dfrac{q}{p^2}$

Demostración.

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} k \ f_X(k) = \sum_{k=1}^{\infty} k \ q^{k-1}p = \sum_{k=1}^{\infty}(k-1+1) \ q^{k-1}p \\ &= \sum_{k=1}^{\infty}(k-1) \ q^{k-1}p + \sum_{k=1}^{\infty} q^{k-1}p \end{aligned}$$

haciendo el cambio de variable $j = k - 1$, obtenemos,

$$= \sum_{j=0}^{\infty} jq^j p + 1 = q\sum_{j=1}^{\infty} jq^{j-1}p + 1 = qE(X) + 1$$

Es decir, $\quad (1-q)E(X) = 1, \quad \therefore E(X) = \dfrac{1}{p}$

$\square$

Generalicemos la distribución Geométrica.

4.5. Distribución Pascal o Binomial negativa

Definición 4.5.1. *La variable aleatoria X que determina, el número de ensayos sucesivos de Bernoulli, independientes y con el mismo parámetro p, hasta que ocurre el r-ésimo éxito E, es conocida como, la variable aleatoria $X \sim P_a(r,p)$ con distribución Pascal o Binomial Negativa con parámetros r,p.*

Teorema 4.5.2 (Distribución Pascal)**.** *La variable aleatoria $X \sim P_a(r,p)$ con distribución Pascal con parámetros r,p, tiene como función de densidad de probabilidad*

$$f_X(k) = \binom{k-1}{r-1} q^{k-r} p^r, \quad \textit{para} \;\; k = r, (r+1), \ldots$$

Teorema 4.5.3. *La Esperanza y la Varianza de la variable aleatoria X, con distribución Pascal, con parámetros r,p, vienen dadas por,*

1. $E(X) = \dfrac{r}{p}$

2. $V(X) = \dfrac{rq}{p^2}$

Demostración.

$$E[X^s] = \sum_{k=r}^{\infty} k^s \; f_X(k) = \sum_{k=r}^{\infty} k^s \binom{k-1}{r-1} q^{k-r} p^r$$

Usando la identidad: $\quad k\binom{k-1}{r-1} = r\binom{k}{r}$

$$E[X^s] = \frac{r}{p} \sum_{k=r}^{\infty} k^{s-1} \binom{k}{r} q^{k-r} p^{r+1}$$

haciendo el cambio de variable $k = j-1$,

$$\begin{aligned} &= \frac{r}{p} \sum_{j=r+1}^{\infty} (j-1)^{s-1} \binom{j-1}{(r+1)-1} q^{j-(r+1)} p^{r+1} \\ &= \frac{r}{p} E[(Y-1)^{s-1}] \end{aligned}$$

donde $Y \sim P_a((r+1),p)$. Con $s=1$, obtenemos

$$E(X) = \frac{r}{p} E[(Y-1)^0] = \frac{r}{p}$$

Y con $s = 2$, como $E(Y) = \dfrac{r+1}{p}$

$$E[X^2] = \frac{r}{p}E[(Y-1)^1] = \frac{r}{p}[E(Y)-1] = \frac{r}{p}\left[\frac{(r+1)}{p} - 1\right]$$

Por lo que,

$$\begin{aligned} V(X) = E[X^2] - (E[X])^2 &= \frac{r}{p}\left[\frac{(r+1)}{p} - 1\right] - \left(\frac{r}{p}\right)^2 \\ &= \frac{r}{p}\left(\frac{(\cancel{r}+1)}{p} - 1 - \frac{\cancel{r}}{p}\right) = \frac{r}{p}\left(\frac{1-p}{p}\right) = \frac{rq}{p^2} \end{aligned}$$

□

Ejemplo 4.5.4. *Un cliente potencial entra a una agencia de automóviles cada hora. La probabilidad de que una vendedora cierre una transacción es 0.10.*

1. *Si ella está dispuesta a continuar trabajando hasta que venda tres carros, ¿cuál es la probabilidad de que trabaje exactamente ocho horas?*

2. *¿Cuántos clientes espera atender para vender el primer auto?*

Sean, $X =$ #de horas, hasta que la vendedora venda tres autos,
$Y =$ #de clientes atendidos, hasta que la vendedora venda el primer auto,
entonces, $X \sim P_a(3, 0.1),\ Y \sim P_a(1, 0.1) = G(0.1),$ con Éxito=Venta de un auto.

1. $P[X = 8] = \dbinom{7}{2}(0.1)^5(0.9)^3 = 0.00015$

2. $E(Y) = \dfrac{1}{0.10} = 10,$ espera atender 10 clientes para lograr su primera venta

4.6. Distribución Poisson

Definición 4.6.1. *Consideremos una fuente de material radiactivo que emite partículas α. Se define a la variable aleatoria con distribución Poisson con parámetro λ_t, como aquella que determina el número de partículas α emitidas X_t durante el intervalo de tiempo $[0, t]$ y que satisface los siguientes cinco axiomas:*

A_1 *Las variables aleatorias X_t y $X[s, s+t]$ para $s > 0$, tienen la misma distribución. Donde, $X[s, s+t] = X[0, s+t] - X[0, s]$, es decir, $X[s, s+t]$ es la variable aleatoria que determina la cantidad de partíclas emitidas durante el intervalo de tiempo $[s, s+t]$.*

A_2 *El número de partículas emitidas durante intervalos de tiempo no sobrepuestos, son variables aleatorias independientes, es decir, para $0 \leq t < t_1 < t_2 < s$, las siguientes variables aleatorias con distribución Poisson son independientes:*

$$X[t, t_1], \;\; X[t_2, s]$$

A_3 *La probabilidad de obtener exactamente una emisión durante un intervalo suficientemente pequeño, es directamente proporcional a la longitud del intervalo.*

A_4 *La probabilidad de obtener al menos dos emisiones en un intervalo suficientemente pequeño es despreciable.*

A_5 *Condición inicial: Para $t = 0$ $\;X_0 = 0$ $\quad$ equivalentemente $\quad P[X_0 = 0] = 1$*

Teorema 4.6.2 (Distribución Poisson)**.** *Bajo los cinco axiomas anteriores la variable aleatoria X_t tiene la siguiente función de distribución de probabilidad*

$$f_{X_t}(k) = P[X_t = k] = e^{-\lambda t}\,\frac{(\lambda t)^k}{k!}, \quad \text{para} \;\; k = 0, 1, 2, \ldots$$

Toda variable que tenga a ésta como fdp. se le conoce como variable con distribución de Poisson con parámetro λt, escribiéndose $X_t \sim P_o(\lambda t)$

Teorema 4.6.3. *La Esperanza y la Varianza de la variable aleatoria X, con distribución Poisson, con parámetro λ, vienen dadas por,*

1. $E(X) = \lambda$
2. $V(X) = \lambda$

Demostración.

$$E[X] = \sum_{k=0}^{\infty} k\; e^{-\lambda}\,\frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty}\,\frac{\lambda^{(k-1)}}{(k-1)!}$$

haciendo el cambio de variable $j = k - 1$, y como $e^{\lambda} = \sum_{j=0}^{\infty}\,\dfrac{\lambda^j}{j!}$

$$E[X] = \lambda e^{-\lambda} \sum_{j=0}^{\infty}\,\frac{\lambda^j}{j!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$E[X^2] = \sum_{k=0}^{\infty} k^2\; e^{-\lambda}\,\frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} k\; e^{-\lambda}\,\frac{\lambda^{(k-1)}}{(k-1)!}$$

haciendo el cambio de variable $j = k - 1$,

$$E[X^2] = \lambda \sum_{j=0}^{\infty} (j+1)\, e^{-\lambda}\, \frac{\lambda^j}{j!} = \lambda E[X+1] = \lambda[E(X)+1] = \lambda(\lambda+1)$$

Por lo que

$$V(X) = E[X^2] - (E[X])^2 = \lambda(\lambda+1) - (\lambda)^2 = \lambda$$

□

Ejemplo 4.6.4. *Suponga que el número de clientes que entran en un banco en una hora es una variable aleatoria Poisson con $P(X = 0) = .05$, ¿cuál es la probabilidad de que en un período de dos horas entre un cliente a este banco?*

Sea $X = \#$ de clientes que entran a este banco en una hora, entonces,

$$0.05 = P[X = 0] = e^{-\lambda}\, \frac{\lambda^0}{0!} \quad \Longrightarrow \quad \lambda = \ln(20) \approx 3$$

por lo que entran al banco en promedio tres clientes por hora.

Sea $Y = \#$ de clientes que entran al banco en dos horas, entonces, $Y \sim P_o(6)$.

Nos piden hallar: $P[Y = 1] = e^{-6} \left(\frac{6}{1!}\right) = 0.01487$

Ejemplo 4.6.5. *El número de buques tanque que llegan cada día a cierta refinería tiene una distribución de Poisson con parámetro λ=2. Las actuales instalaciones portuarias pueden despachar tres buques al día. Si más de tres buques tanque llegan en un día, los restantes deben enviarse a otro puerto.*

1. *En un día determinado, ¿cuál es la probabilidad de tener que hacer salir buques tanque?*
2. *¿En cuánto deben aumentar las instalaciones actuales para permitir la atención a todos los buques tanque aproximadamente el 90 % de los días?*
3. *¿Cuál es el número más probable de buques tanque que llegan diariamente?*

4. *¿Cuál es el número esperado de buques tanque que atienden diariamente?*

5. *¿Cuál es el número esperado de buques tanque devueltos diariamente?*

Sean $X =\#$ de buques tanque que llegan cada día a la refinería,
$Y =\#$ de buques tanque que son atendidos diariamente en la refinería,
$Z =\#$ de buques tanque que son devueltos diariamente en la refinería,
entonces: $X \sim P_o(\lambda = 2)$, por lo que $E(X) = 2$.

1. Se harán salir buques tanque si llegan más de tres.

$$P[X \geq 4] = 1 - P[X \leq 3] = 1 - e^{-2}\left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}\right) = 1 - 0.8571 = 0.1429$$

2. $P[X \leq n] = 0.90$ nos indica que con probabilidad 0.90 llegan a lo más n buques por día, en otras palabras, el 90 % de las veces (días) llegarán a lo más n buques.

$$P[X \leq 3] = 0.8571, \quad \text{por 1.}$$
$$P[X \leq 4] = e^{-2}\left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!}\right) = 0.94731$$

el 94.7 % de los días llegarán a lo más 4 buques.

Para garantizar, que sean atendidos todos los buques que llegan a la refinería el 90 % de los días, hay que extender las instalaciones portuarias para que puedan despachar cuatro buques por día.

3. Nos piden hallar n, con $P[X = n]$ máximo.

$$P[X = 0] = e^{-2} \qquad P[X = 1] = 2e^{-2}$$
$$P[X = 2] = 2e^{-2} \qquad P[X = 3] = \tfrac{4}{3}e^{-2}$$

Para $n = 4, 5, \ldots;\ k = 1, 2;\quad P[X = n] \ < \ P[X \geq 4] \ < \ P[X = k]$.

El número más probable de buques que llegan diariamente es de $n = 1, 2$, con,

$$P[X = n] = 2e^{-2} = 0.270670566 \quad \text{máximo.}$$

4. Se sabe que para $k = 0, 1, 2;$ k buques serán atendidos si y solo si k buques llegan al día, y tres buque serán atendidos si y solo si tres o más buques llegan por día, por lo que,

$$[Y = 0] = [X = 0], \quad [Y = 1] = [X = 1], \quad [Y = 2] = [X = 2], \quad [Y = 3] = [X \geq 3]^c$$

$$\begin{aligned} E(Y) =& \cancel{0P[X = 0]} + 1P[X = 1] + 2P[X = 2] + 3(1 - P[X \leq 2]) \\ =& 2e^{-2} + 4e^{-2} + 3(1 - 5e^{-2}) = 3 - 9e^{-2} = 1.782 \end{aligned}$$

5. De la relación: $Z = X - Y$, obtenemos,

$$E(Z) = E(X) - E(Y) = 2 - 1.782 = 0.218$$

El número esperado de buques tanque devueltos diariamente es de 0.218

Nota 4.6.6. *Para variables aleatorias de uso común los valores de la función de distribución $F_X(x) = P[X \leq x]$ y de su inversa $F_X^{-1}(x)$ se podrán obtener con el paquete estadístico R con las siguientes instrucciones.*

Instrucción R		*Distribución*
$F_X(x)$	$F_X^{-1}(x)$	*Discreta*
$>$ `pbinom(x,n,p)`	$>$ `qbinom(x,n,p)`	$B(n, p)$
$>$ `phyper(x,r,N-r,n)`	$>$ `qhyper(x,r,N-r,n)`	$H(r, n, N)$
$>$ `pnbinom(x,r,p)`	$>$ `qnbinom(x,r,p)`	$P_a(r, p)$
$>$ `ppois(x,`λ`)`	$>$ `qpois(x,`λ`)`	$P_o(\lambda)$

Instrucción R		*Distribución*
$F_X(x)$	$F_X^{-1}(x)$	*Continua*
$>$ `punif(x,a,b)`	$>$ `qunif(x,a,b)`	$U(a, b)$
$>$ `pnorm(x,`μ,σ`)`	$>$ `qnorm(x,`μ,σ`)`	$N(\mu, \sigma)$
$>$ `pexp(x,`λ`)`	$>$ `qexp(x,`λ`)`	$E(\lambda)$
$>$ `pgamma(x,n,`λ`)`	$>$ `qgamma(x,n,`λ`)`	$\Gamma(n, \lambda)$

Ejemplo 4.6.7.

Sea $X \sim B(5, 0.4)$ obtengamos $P[X = 5] = f(5)$, $P[X \leq 4] = F_X(4)$, $F_X^{-1}(0.03856)$. Con fórmula

$$\begin{aligned} & f(5) = P[X = 5] = \binom{5}{5}(0.4)^5 = 0.01024 \\ & F_X(4) = P[X \leq 4] = 1 - f(5) = 1 - 0.01024 = 0.98976 \\ & \therefore \quad F_X^{-1}(0.98976) = 4 \end{aligned}$$

Con paquete estadístico R

```
> pbinom(5, 5, 0.4) - pbinom(4, 5, 0.4)
[1] 0.01024
> pbinom(4, 5, 0.4)
[1] 0.98976
> qbinom(0.98976, 5, 0.4)
[1] 4
```

Por lo tanto

$$P[X = 5] = f(5) = F_X(5) - F_X(4) = 0.01024$$
$$P[X \leq 4] = F_X(4) = 0.98976$$
$$F_X^{-1}(0.98976) = 4$$

Resumen de distribuciones discretas de uso común					
Distribución	Variable aleatoria	Parám.	Función de distribución de probabilidades	Media	Varianza
1. Binomial $X \sim B(n,p)$	X =Números de éxitos en n ensayos independientes de bernoulli	$n,\ p$	$f(k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, \ldots, n$	np	npq
2. Geométrica $X \sim G(p)$	X =Números de ensayos independientes de bernoulli asta que ocurre el primer éxito	p	$f(k) = q^{k-1}p, \quad k = 1, 2 \ldots$	$\frac{1}{p}$	$\frac{q}{p^2}$
3. Pascal $X \sim P_a(r,p)$	X =Números de ensayos independientes de bernoulli asta que ocurre el r-ésimo éxito	$r,\ p$	$f(k) = \binom{k-1}{r-1} q^{k-r} p^r, \quad k = r, (r+1), \ldots$	$\frac{r}{p}$	$\frac{rq}{p^2}$
4. Hipergeom. $X \sim H(r,n,N)$	X =Números de éxitos en n extracciones sin reemplazo	$r,\ n,\ N$	$f(k) = \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}}, \quad k = 0, \ldots \text{mín}\{r, n\}$	np	$npq \cdot \left(\frac{N-n}{N-1}\right)$, donde $p = \frac{r}{N}$
5. Poisson $X \sim P(\lambda)$	X =Números de éxitos en un intervalo	λ	$f(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2 \ldots$	λ	λ
Relaciones entre distribuciones					
Si $Y \sim P_a(r,p)$ y $X \sim B(n,p)$, entonces: $P(Y \leq n) = P(X \geq r),\ P(Y > n) = P(X < r)$			Si $X \sim H(r,n,N)$ y $\frac{n}{N} \leq 0.1$, entonces $X \sim B\left(n, \frac{r}{N}\right)$		
Para n suficientemente grande con $p \approx 0$, si $X \sim B(n,p)$, entonces $X \approx P(np)$.					

4.7. Ejercicios

1. Supóngase que la máquina 1 produce (diariamente) el doble de artículos que la máquina 2. Sin embargo, cerca del 4 % de los artículos de la máquina 1 tiende a ser defectuosos, mientras que la máquina 2 sólo produce alrededor de 2 % defectuosos. Supongamos que se combina la producción diaria de las dos máquinas. Se toma una muestra aleatoria de diez del resultado combinado. ¿Cuál es la probabilidad de que esta muestra contenga dos defectuosos?

2. Una fábrica produce diariamente diez recipientes de vidrio. Se puede suponer que hay una probabilidad constante $p = 0.1$ de producir uno defectuoso. Antes de que estos recipientes se almacenen son inspeccionados y a los defectuosos se les aparta. Supongamos que hay una probabilidad constante $r = 0.1$ de que un recipiente defectuoso sea mal clasificado. Sea X igual al número de recipientes clasificados como defectuosos al término de un día de producción. (Suponemos que todos los recipientes que se fabrican en un día se inspeccionan ese mismo día y que la probabilidad de que un recipiente no defectuoso sea mal clasificado es cero ya que en caso contrario se le correría al personal de inspección.)
   Calcular $P[X = 3]$ y $P(X > 3)$,

3. Una compañía de seguros ha descubierto que sólo alrededor del 0.1 % de la población tiene cierto tipo d e accidente cada año. Si los 10 O00 asegurados fueran seleccionados aleatoriamente en la población, ¿cuál será la probabilidad de que no más de 5 de estos clientes tenga un accidente de ese tipo el próximo año?

4. Una fuente radiactiva se observa durante 7 intervalos cada uno de 10 segundos de duración y se cuenta el número de partículas emitidas durante cada periodo. Suponer que el número de partículas emitidas, digamos X , durante cada periodo observado tiene una distribución d e Poisson con parámetro 5.0 (es decir, las partículas se emiten a razón de 0.5 partículas por segundo).

   a) ¿Cuál es la probabilidad de que en cada uno de los 7 intervalos de tiempo, se emitan 4 o más partículas?

   b) ¿Cuál es la probabilidad de que al menos en uno de los 7 intervalos de tiempo se emitan 4 o más partículas?

5. La probabilidad de que el lanzamiento de un cohete sea exitoso es igual a 0.8. Supóngase que se hacen ensayos hasta que ocurren 3 lanzamientos esitosos.

   a) ¿Cuál es la probabilidad de que sean necesarios 6 intentos?

   b) ¿Cuál es la probabilidad de que sean necesarios menos de 6 intentos?

6. En la situación descrita en el problema anterior suponer que los ensayos de lanzamiento se hacen hasta que ocurren tres lanzamientos consecutivos exitosos.

   a) ¿Cuál es la probabilidad de que sean necesarios 6 intentos?

   b) ¿Cuál es la probabilidad de que sean necesarios menos de 6 intentos?

Capítulo 5

Variables aleatorias continuas de uso común

En esta sección estudiamos a las variables aleatorias continuas que más se usan, de éstas la de mayor úso sin menospreciar a las otras, es la variable aleatoria normal, en parte ésto se debe al famoso teorema del límite central (ver [2], [3]).

5.1. Distribución Uniforme

Definición 5.1.1. *Diremos que la variable aleatoria continua X tiene distribución Uniforme en el intervalo $[a, b]$, escribiéndose $X \sim U[a, b]$, si tiene la siguiente función de densidad*

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & si \;\; a \leq x \leq b, \\ 0, & en \; otro \; caso. \end{cases}$$

Teorema 5.1.2. *Sea X una variable aletoria con distribución Uniforme en el intervalo $[a, b]$; entonces:*

1. $F_X(x) = \begin{cases} \dfrac{x-a}{b-a}, & si \;\; a \leq x \leq b, \\ 0, & en \; otro \; caso. \end{cases}$

2. $E(X) = \dfrac{a+b}{2}$

3. $V(X) = \dfrac{(b-a)^2}{12}$

Demostración.

1. Obtengamos la función de distribución F_X

$$F_X(x) = P[X \leq x] = \int_{-\infty}^{x} f_X(y)\, dy = \begin{cases} 0, & \text{si } x < a, \\ \int_a^x \frac{1}{b-a}\, dy = \frac{x-a}{b-a}, & \text{si } a \leq x \leq b, \\ 1 & \text{si } b < x. \end{cases}$$

$$E(X) = \int_{\mathbb{R}} x f_X(x)\, dx = \frac{1}{b-a}\int_a^b x\, dx = \frac{1}{2(b-a)} x^2|_b^a = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

$$E(X^2) = \int_{\mathbb{R}} x^2 f_X(x) dx = \frac{1}{b-a}\int_a^b x^2 dx = \frac{1}{3(b-a)} x^3|_b^a = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ba+a^2}{3}$$

$$V(X) = E(X^2) - (E(X))^2 = \frac{b^2+ba+a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$$

$\square$

La Distribución Uniforme se presenta cuando elegimos al azar un punto del intervalo $[a,b]$.

Ejemplo 5.1.3. *Se escoge al azar un punto x para dividir un segmento de longitud L. ¿Cuál es la probabilidad de que la razón del segmento más corto en relación con el más largo sea menor que $1/4$?*

Sea X la variable aleatoria que determina el punto x elegido al azar en el segmento $[0,L]$. Nos piden hallar $P(A)$, donde $A = \left[\frac{X}{L-X} < \frac{1}{4}\right] \cup \left[\frac{L-X}{X} < \frac{1}{4}\right]$. Como $X \sim U[0,L]$,

$$\begin{aligned} P(A) =& P\left[\frac{X}{L-X} < \frac{1}{4}\right] + P\left[\frac{L-X}{X} < \frac{1}{4}\right] = P\left[4X < L - X\right] + P\left[4L - 4X < X\right] \\ =& P\left[X < \frac{L}{5}\right] + P\left[X > \frac{4L}{5}\right] = F_x(L/5) + (1 - F_x(4L/5)) \\ =& \frac{L/5 - 0}{L - 0} + \left(1 - \frac{4L/5 - 0}{L - 0}\right) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5} \end{aligned}$$

5.2. Distribución Normal

Definición 5.2.1. *La variable aleatoria continua X que toma valores reales tiene una distribución Normal o Gaussiana con parámetros $\mu \in \mathbb{R}$, y $\sigma^2 > 0$, escribiéndose $X \sim N(\mu, \sigma^2)$; si su función de densidad está dada por*

$$f_X(x) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Teorema 5.2.2. *Si X una variable aletoria con distribución Normal con parámetros μ, σ^2; entonces:*

1. $E(X) = \mu$.

2. $V(X) = \sigma^2$.

Demostración.

$$\begin{aligned}
E(X) =& \int_{\mathbb{R}} x f_X(x)\ dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ dx/\sigma \\
=& \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (\sigma z + \mu)\ e^{-z^2/2}\ dz \\
=& \cancelto{0}{\frac{\sigma}{\sqrt{2\pi}} \int_{\mathbb{R}} z\ e^{-z^2/2}\ dz} + \mu \cancelto{1}{\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}\ dz} \\
=& \mu
\end{aligned}$$

Por ser la penúltima, una integral de una función impar; y la última, una integral de una función de densidad f_Y de una variable aleatoria $Y \sim N(0,1)$.

Nuevamente, utilizando el cambio de variable $z = (x-\mu)/\sigma$, obtenemos

$$\begin{aligned}
E(X^2) =& \int_{\mathbb{R}} x^2 f_X(x)\ dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^2\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ dx/\sigma \\
=& \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (\sigma z + \mu)^2\ e^{-z^2/2}\ dz \\
=& \frac{\sigma^2}{\sqrt{2\pi}} \int_{\mathbb{R}} z^2\ e^{-z^2/2}\ dz + \cancelto{0}{\frac{2\mu\sigma}{\sqrt{2\pi}} \int_{\mathbb{R}} z\ e^{-z^2/2}\ dz} + \mu^2 \cancelto{1}{\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}\ dz}.
\end{aligned}$$

En el primer término integrando por partes, haciendo $u = z$ y $dv = z\ e^{-z^2/2}\ dz$; obtenemos

$$\begin{aligned}
E(X^2) =& \cancelto{0}{\left[\frac{-\sigma^2\ z\ e^{-z^2/2}}{\sqrt{2\pi}}\right]\Bigg|_{-\infty}^{+\infty}} + \sigma^2 \cancelto{1}{\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}\ dz} + \mu^2 = \sigma^2 + \mu^2 \\
\therefore \qquad & V(X) = E(X^2) - (E(X))^2 = (\sigma^2 + \mu^2) - \mu^2 = \sigma^2.
\end{aligned}$$

□

A continuación enunciamos tres Teoremas, sin demostración.

Teorema 5.2.3. *Sea X variable aleatoria con distribución $N(\mu, \sigma^2)$, entonces la variable $Y = aX + b$ tiene distribución $N(a\mu + b, a^2\sigma^2)$.*

Corolario 5.2.4. *Si la variable aleatoria X tiene distribución $N(\mu, \sigma^2)$, entonces la variable $Z = \dfrac{X-\mu}{\sigma}$ tiene distribución $N(0,1)$. La cual es conocida como variable aleatoria Normal estándar y su función de distribución F_Z será denotada por la letra Φ.*

Ejemplo 5.2.5. *En una linea de producción, el diámetro de un cable eléctrico está distribuido normalmente con promedio 0.8 pulgadas y desviación estandar 0.02 pulgadas, supóngase que los cables se consideran defectuosos si su diámetro se diferencia de su promedio en más de 0.025 pulgadas. ¿Cuál es la probabilidad de obtener un cable defectuoso?*

Sea D la variable aleatoria que determina el diámetro de un cable eléctrico producido, es decir $D \sim N(0.8, .0004)$. Como un cable será defectuosos si $|D-\mu| > 0.025$, entonces,

$$\begin{aligned} P[|D-\mu| > 0.025] &= P\left[\left|\frac{D-\mu}{\sigma}\right| > \frac{0.025}{0.02}\right] = P(Z| > 1.25) = 1 - P(Z| \leq 1.25) \\ &= 1 - P(-1.25 \leq Z \leq 1.25) \end{aligned}$$

Por Tabla 1. de la Normal estandar

$$P[|D-\mu| > 0.025] = 1 - (\Phi(1.25) - \Phi(-1.25)) = 1 - (0.8944 - 0.1056) = 0.2112$$

Ejemplo 5.2.6. *Suponga que las calificaciones de las pruebas de admisión a una universidad tienen distribución normal, con media 450 y 100 de desviación estándar.*

1. *Si esta universidad no admite a quienes tengan menos de 480 de calificación, ¿qué porcentaje de las personas que presentan el examen calificarían para ingresar a dicha universidad?*

2. *¿Cuál sería la calificación mínima C, para estar dentro del 5 % de los mejores calificados?*

Sea X la v.a. que determina la calificación en dicha prueba, entonces, $X \sim N(450, 100^2)$.

1. $$\begin{aligned} P[X \geq 480] =& P\left[Z \geq \frac{480-450}{100}\right] = P[Z \geq 0.3] = 1 - P[Z \leq 0.3] = 1 - \Phi(0.3) \\ =& 1 - 0.6179 = 0.3821 \end{aligned}$$

El 38.21 % de las personas calificarían para ingresar a dicha universidad.

2. $$\begin{aligned} P[X \geq C] =& 0.05 \\ \implies& P\left[Z \geq \frac{C-450}{100}\right] = 1 - \Phi\left(\frac{C-450}{100}\right) = 0.05 \\ \implies& \Phi\left(\frac{C-450}{100}\right) = 0.95 \\ \implies& \frac{C-450}{100} = \Phi^{-1}(0.95) = 1.645 \\ \implies& C = (100)(1.645) + 450 = 614.5 \end{aligned}$$

La calificación mínima sería $C = 614.5$

Teorema 5.2.7 (Propiedad reproductiva de la distribución Normal)**.** *La suma de variables aleatorias X_i independientes e idénticamente distribuidas $X_i \sim N(\mu_i, \sigma_i^2)$; es una variable Y, normalmente distribuida, más aún, $Y \sim N(\sum \mu_i, \sum \sigma_i^2)$.*

Teorema 5.2.8 (Teorema del límite central)**.** *Sean $X_1, X_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas con media μ y varianza σ^2 finitas. Entonces para n suficientemente grande la variable aleatoria $X_1 + X_2 + \cdots + X_n$ tiene aproximadamente una distribución normal con media $n\mu$ y varianza $n\sigma^2$ en el sentido de que para toda $x \in \mathbb{R}$*

$$\lim_{n\to\infty} P(X_1 + X_2 + \cdots + X_n \leq x) = P(X_N \leq x), \qquad X_N \sim N(n\mu, n\sigma^2)$$

5.2.1. Aproximación Normal a la distribución Binomial

Teorema 5.2.9. *Sea $X \sim B(n,p)$ entonces, para n suficientemente grande la variable aleatoria X Binomial se puede aproximar por una variable X_N normal, es decir*

$$P(X \leq x) \approx P(X_N \leq x)$$

donde $X_N \sim N(np, npq)$.

Demostración. Se sigue de Nota 4.2.2 y Teorema 5.2.8. □

Nota 5.2.10.

1. *Cuando se aproxima una distribución discreta con una distribución continua, como es en este caso, hay que considerar un factor de corrección*

$$P[a \leq X \leq b] \approx P[a - \tfrac{1}{2} \leq X_N \leq b + \tfrac{1}{2}]$$

2. *Para obtener una buena aproximación se recomienda*

$$np \geq 5, \quad nq \geq 5.$$

Ejemplo 5.2.11. *Para determinar la efectividad de una dieta para reducir la cantidad de colesterol en el torrente sanguíneo, se somete a esta dieta a 100 personas en un período de tiempo suficiente, y finalmente se compara los dos niveles de colesterol de cada uno de los pacientes (inicio y término del experimento). Si un nutriólogo decide respaldar la dieta si al menos el 65 % de las personas logra reducir su nivel de colesterol. ¿Cuál es la probabilidad de que se respalde la dieta cuando realmente ésta no reduce el nivel de colesterol?*

Sea $X = \#$ de personas sometidas a la dieta que redujeron su nivel de colesterol, entonces

$$X \sim B_{in}(100, 0.5), \quad \text{con Éxito=paciente con reducción de su nivel de colesterol}$$

Como la dieta no tiene nada que ver con la reducción del nivel de colesterol, ésta reducción se deberá al azar, por lo que, $p = 0.5$. Nos piden hallar $P[X \geq 65]$, utilizando R

$> 1 - \texttt{pbinom}(64, 100, .5)$
[1] 0.001758821
$\therefore \quad P[X \geq 65] = 0.001758821$

Otra forma de resolver este problema es aproximando con la distribución normal, ya que $np \geq 5$ y $nq \geq 5$. Sea $X_N \sim N(50, 25)$.

$$P[X \geq 65] \approx P[X_N \geq 64.5] = P[Z \geq \frac{64.5 - 50}{5}] = P[Z \geq 2.9] = 1 - \Phi(2.9) = 1 - 0.9981$$

$\therefore \quad P[X \geq 65] \approx 0.0019$

5.3. Distribución Gamma

Definición 5.3.1. *La variable aleatoria continua X que toma valores no negativos tiene una distribución Gamma con parámetros $r > 0$ y $\lambda > 0$, escribiéndose $X \sim \Gamma(r, \lambda)$; si su función de densidad está dada por*

$$f_X(x) = \begin{cases} \dfrac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}, & \text{si } \; x \geq 0, \\ 0, & \text{en otro caso.} \end{cases}$$

donde Γ es la conocida función Gamma, es decir, para $r > 0$

$$\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}\; dx.$$

Nota 5.3.2. *1.* $\Gamma(1) = \displaystyle\int_0^\infty e^{-x}\; dx = \lim_{M\to\infty} -e^{-x}\Big|_0^M = 1$

2. *Cambiando de variable, tomando* $x = \frac{1}{2}z^2,\; dx = z\; dz,\; x^{-1/2} = \frac{\sqrt{2}}{z}$

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty \tfrac{\sqrt{2}}{z}e^{-\frac{z^2}{2}}z\; dz = \sqrt{2}\,\sqrt{2\pi}\int_0^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\; dz = 2\sqrt{\pi}P[Z \geq 0] = \sqrt{\pi}$$

3. *Integrando por partes, con* $u = x^{r-1},\; dv = e^{-x}dx$. *Si* $r > 0$

$$\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}\; dx = x^{r-1}e^{-x} + (r-1)\int_x^\infty x^{(r-1)-1}e^{-x}\; dx = (r-1)\;\Gamma(r-1))$$

Por lo tanto, denotando a $\Gamma(1) = 1$ *por* $0!$ *si* $n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$

Teorema 5.3.3. *Sea X una variable aletoria con distribución Gamma con parámetros r, λ*

1. $E(X) = \dfrac{r}{\lambda}$.
2. $V(X) = \dfrac{r}{\lambda^2}$.

Demostración.

1. Integrando por partes, haciendo $u = \dfrac{(\lambda x)^r}{\lambda\Gamma(r)}$ y $dv = \lambda e^{-\lambda x}\; dx$; obtenemos

$$E(X) = \int_{\mathbb{R}} x f_X(x)\; dx = \int_0^\infty x\;\frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}\; dx$$

$$= \cancelto{0}{\left[-\frac{(\lambda x)^r}{\lambda\Gamma(r)}e^{-\lambda x}\right]\Bigg|_0^\infty} + \int_0^\infty \frac{r(\lambda x)^{r-1}}{\Gamma(r)}e^{-\lambda x}\; dx = \frac{r}{\lambda}\cancelto{1}{\int_{\mathbb{R}} f_X(x)\; dx} = \frac{r}{\lambda}$$

2. Integrando por partes, haciendo $u = \dfrac{(\lambda x)^{r+1}}{\lambda^2\Gamma(r)}$ y $dv = \lambda e^{-\lambda x}\ dx$; obtenemos

$$\begin{aligned}
E(X^2) &= \int_{\mathbb{R}} x^2 f_X(x)\ dx = \int_0^\infty x^2\ \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}\ dx \\
&= \cancel{\left[-\frac{(\lambda x)^{r+1}}{\lambda^2\Gamma(r)}e^{-\lambda x}\right]\Bigg|_0^\infty} \rightarrow 0 + \int_0^\infty \frac{(r+1)(\lambda x)^r}{\lambda\Gamma(r)}e^{-\lambda x}\ dx \\
&= \frac{(r+1)}{\lambda}\int_{\mathbb{R}} x f_X(x)\ dx = \frac{(r+1)}{\lambda}\ E(X) = \frac{r(r+1)}{\lambda^2} \\
\therefore \quad & V(X) = E(X^2) - (E(X))^2 = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}
\end{aligned}$$

□

Casos particulares de la distribución Gamma

5.3.1. Distribución Exponencial

La distribución Exponencial $Exp(\lambda)$ es el caso especial de una distribución gamma $\Gamma(\lambda, r)$, con parámetro $r = 1$.

Definición 5.3.4. *La variable aleatoria continua X que toma valores no negativos tiene una distribución Exponencial con parámetro $\lambda > 0$, si su función de densidad está dada por*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & si\ \ x \geq 0, \\ 0, & en\ otro\ caso. \end{cases}$$

Teorema 5.3.5. *Si X es una variable aletoria con distribución Exponencial con parámetro λ; entonces,*

1. $F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & si\ \ x \geq 0, \\ 0, & en\ otro\ caso. \end{cases}$
2. $P[X > s + t | X > s] = P[X > t]$ *(Propiedad de no tener memoria).*
3. $E(X) = \dfrac{1}{\lambda}$.
4. $V(X) = \dfrac{1}{\lambda^2}$.

Demostración.

1. Obtengamos la función de distribución F_X.

$$F_X(x) = P[X \leq x] = \int_{-\infty}^x f_X(y)\ dy = \begin{cases} 0, & \text{si}\ \ x < 0, \\ \int_0^x \lambda e^{-\lambda y}\ dy = 1 - e^{-\lambda x}, & \text{si}\ \ x \geq 0. \end{cases}$$

2. La distribución Exponencial tiene la propiedad de "no tener memoria".

$$P[X > s+t | X > s] = \frac{P[X > s+t]}{P[X > s]} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P[X > t].$$

3. y 4. se siguen de Teorema5.3.3 con $r = 1$.

□

5.3.2. Distribución Erlang

La Distribución Erlang es el caso particular de una distribución Gamma $\Gamma(\lambda, n)$ con parámetro $n \in \mathbb{N}$. En este caso la distribución Gamma está relacionada con la distribución Poisson y n variables aleatorias Exponenciales con parámetro λ.

Definición 5.3.6. *La variable aleatoria continua X que toma valores no negativos tiene una distribución Erlang con parámetros $\lambda > 0, n$, si su función de densidad está dada por*

$$f_X(x) = \begin{cases} \dfrac{\lambda}{(n-1)!}(\lambda x)^{n-1} e^{-\lambda x}, & si \;\; x \geq 0, \\ 0, & en\ otro\ caso. \end{cases}$$

Teorema 5.3.7. *Si $X \sim \Gamma(\lambda, n)$ con $n \in N$; entonces:*

1. $F_X(x) = \begin{cases} 0, & si \;\; x \leq 0, \\ 1 - F_Y(n-1), & si \;\; x > 0. \end{cases};$ *donde $Y \sim P_o(\lambda x)$.*

2. $E(X) = \dfrac{n}{\lambda}$.

3. $V(X) = \dfrac{n}{\lambda^2}$.

Demostración.

1. Para $n \geq 2$, integrando por partes, haciendo $u = \dfrac{(\lambda y)^{n-1}}{(n-1)!}$ y $dv = \lambda e^{-\lambda y}\, dy$, para $x > 0$

$$\begin{aligned} P[X > x] &= \int_x^\infty \frac{\lambda}{(n-1)!}(\lambda y)^{n-1} e^{-\lambda y}\, dy \\ &= \left[-\frac{(\lambda y)^{n-1}}{(n-1)!} e^{-\lambda y} \right]\Bigg|_x^\infty + \int_x^\infty \frac{\lambda (\lambda y)^{n-2}}{(n-2)!} e^{-\lambda y}\, dy = \frac{(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} + P[X_{n-1} > x] \end{aligned}$$

donde $X_{n-1} \sim \Gamma(\lambda, n-1)$, por Teorema 5.3.5 $P[X_1 > x] = e^{-\lambda x}$, entonces,

Para $x > 0$, $n \geq 2$,

$$P[X > x] = e^{-\lambda x}\left[\frac{(\lambda x)^{n-1}}{(n-1)!} + \cdots + \frac{(\lambda x)^2}{2!} + \frac{\lambda x}{1!} + 1\right] = P[Y \leq n-1] = F_Y(n-1)$$

donde $Y \sim P_o(\lambda x)$

$$\therefore \quad F_X(x) = \begin{cases} 0, & \text{si} \;\; x \leq 0, \\ 1 - P[X > x] = 1 - F_Y(n-1), & \text{si} \;\; x > 0. \end{cases}$$

2. y 3. se siguen de Teorema5.3.3 con $r = n$.

$\square$

A continuación enunciamos dos teoremas sin demostración.

Teorema 5.3.8. *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con la distribución exponencial con parámetro λ, entonces la variable $T_n = X_1 + X_2 + \cdots + X_n$ tiene distribución Gamma con parámetros λ, n.*

Teorema 5.3.9. *En todo proceso de Poisson X_t con parámetro λt, el tiempo aleatorio X_1 en que se emite la primera partícula α y los tiempos aleatorios X_{n+1}, $n \in \mathbb{N}$ entre la $(n-1)$-ésima y la n-ésima partícula α emitida son variables independientes con distribución Exponencial con el mismo parámetro λ. Por lo que el tiempo aleatorio T_n en que se emite la n-ésima partícula α tiene distribución Gamma con parámetros λ, n, es decir:*

$$P[T_n = X_1 + X_2 + \cdots + X_n \leq t] = P[X_t \geq n] = 1 - \sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Ejemplo 5.3.10. *Un proceso de fabricación produce en promedio un artículo defectuoso entre 300 fabricados. ¿Cuál es la probabilidad de que aparezca el tercer artículo defectuoso:*

1. *antes de que sean producidos 1 000 artículos?*
2. *cuando se produce el 1 000-ésimo artículo?*
3. *después de que se produzca el 1 000-ésimo artículo?*

Tómese como ensayos independientes de Bernoulli con parámetro $p = 1/300$ a las observaciones de los artículos producidos.

Sea $X =$ Número necesario de artículos producidos para obtener el tercer artículo defectuoso, Obsérvese que $X \sim P_a(3, 1/300)$, y nos piden hallar,

1. $P[X < 1\,000]$, 2. $P[X = 1\,000]$, 3. $P[X > 1\,000]$

Utilizando el paquete estadídtico R, obtenemos los sigs. valores $F_X(x)$

> $>$ `pnbinom(999,3,1/300)`
> [1] 0.6489855
> $>$ `pnbinom(1000,3,1/300)`
> [1] 0.6496445

Por lo que

1. $P[X < 1000] = F_X(999) = 0.6489855$
2. $P[X = 1000] = F_X(1000) - F_X(999) = 0.6496445 - 0.6489855 = 0.000659$
3. $P[X > 1000] = 1 - F_X(1000) = 1 - 0.6496445 = 0.3503555$

Segundo método Resolvamos el ejercicio utilizando la distribución Poisson.

Sea $Y =$ Número de artículos defectuosos en t artículos producidos, entonces $Y \sim B(t, 1/300)$, siendo $p = 1/300$ muy pequeño, entonces Y se aproxima a $X_t \sim P_o((1/300)t)$.
Por Teorema 5.3.9 $\quad X \sim \Gamma(3, 1/300)$ aproximadamente

1. $P[X < 1000] = P[T_3 \leq 1000] = 1 - \left[e^{-10/3}\left(1 + \frac{10/3}{1!} + \frac{(10/3)^2}{2!}\right)\right] = 0.64326$
2. $P[X = 1000] = P[T_3 = 1000] = 0$
3. $P[X > 1000] = 1 - P[X \leq 1000] = 0.35674$

En este segundo método resolvimos el problema de forma aproximada, pero sin la ayuda de un paquete estadístico.

Ejemplo 5.3.11. *Supóngase que el número de accidentes en un fábrica se puede expresar por un proceso de Poisson con un promedio de 2 accidentes por semana. ¿Cuál es la probabilidad de que:*

1. *el tiempo entre un accidente y el siguiente sea mayor de 3 días?*

2. *el tiempo de un accidente al tercero sea mayor de una semana?*

Sean $X_t \sim P(\lambda t)$, $T_1 \sim Exp(\lambda) = \Gamma(1, \lambda)$, $T_2 \sim \Gamma(2, \lambda)$, con $\lambda = 2/7$ y t en días, donde,

$X_t = \#$número de accidentes en la fábrica.

T_1 =Tiempo entre un accidente y el siguiente.

T_2 =Tiempo entre un accidente al tercero.

$$P[T_1 > 3] = e^{-3\lambda} = e^{-6/7} = 0.424373$$

Por Teorema 5.3.9 $P[T_2 > 7] = F_Y(n-1=1)$, con $Y \sim P_{ois}(7\lambda = 2)$, por lo que

$$P[T_2 > 7] = e^{-7\lambda}\left(1 + \frac{7\lambda}{1!}\right) = e^{-2}(1+2) = 0.406006$$

En los dos ejemplos anteriores podemos observar como están relacionadas las tres distribuciones: $E(\lambda), P_o(\lambda), \Gamma(n, \lambda)$.

5.3.3. Distribución Ji-cuadrada

El caso especial de la distribución Gamma $\Gamma(1/2, n/2)$ es conocida como: la Distribución χ^2_n Ji-cuadrada con $n \in \mathbb{N}$ grados de libertad. Así

Definición 5.3.12. *La variable aleatoria continua X con valores no negativos, tiene una distribución Ji-cuadarad con n grados de libertad si su función de densidad está dada por*

$$f_X(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} x^{(n/2)-1} e^{-x/2}, & \text{si } x \geq 0, \\ 0, & \text{en otro caso.} \end{cases}$$

Por Teorema5.3.3, se sigue

Corolario 5.3.13. *Si $X \sim \chi^2_n$, entonces, $E(X) = n, \quad V(X) = 2n$*

Resumen de distribuciones continuas importantes					
Distribución	Parámetros	Función de densidad de probabilidades	F.d.a	Media	Varianza
1. Uniforme	$X \sim U(a,b)$	$f(x) = \frac{1}{b-a} I_{[a,b]}$	$F(x) = \frac{x-a}{b-a} I_{[a,b]} + I_{(b,\infty)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
2. Exponencial	$X \sim Exp(\lambda)$	$f(x) = \lambda e^{-\lambda x} I_{(0,\infty)}$	$F(x) = \left(1 - e^{-\lambda x}\right) I_{(0,\infty)}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
3. Gamma	$X \sim G(n,\lambda)$	$f(x) = \frac{\lambda}{(n-1)!} (\lambda x)^{n-1} e^{-\lambda x} I_{(0,\infty)}$	$F(x) = \left(1 - \sum_{k=0}^{n-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!}\right) I_{(0,\infty)}$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
4.Ji-cuadrada	$X \sim \chi^2_n$	$f(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}} I_{(0,\infty)}$	Tablas	n	$2n$
2. Normal	$X \sim N(\mu,\sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	Tablas	μ	σ^2
Relaciones entre distribuciones					
Para n suficientemente grande con $np \geq 5$ y $nq \geq 5$, si X tiene distribución binomial (n,p), entonces $X \sim N(np, npq)$.					

5.4. Ejercicios

1. Supóngase que X está distribuida uniformemente en $[-a, a]$, donde $a > 0$. Cada vez que sea posible, determinar a de modo que se satisfaga lo siguiente.

   a) $P[X > 1] = 1/3$ *b*) $P[X > 1] = 1/2$ *c*) $P[|X| < 1] = P[|X| > 1]$
   d) $P[X < 1/2] = 0.3$ *e*) $P[X < 1/2] = 0.7$

2. Si la variable aleatoria K está distribuida uniformemente en $[0, 5]$, ¿cuál es la probabilidad de que las raíces de la ecuación $4x^2 + 4xK + K + 2 = 0$ sean reales?

3. Se sabe que la lluvia anual que cae en cierta región es una variable aleatoria distribuida normalmente con media igual a 29.5 pulgadas y desviación estándar 2.5 pulgadas. ¿Cuántas pulgadas de lluvia (anuales) caen en exceso alrededor del 5 % de las veces?

4. Suponga que el tiempo transcurrido hasta la falla de un horno de microondas se distribuye exponencialmente con media de tres años. Una compañia ofrece garantía por el primer año de uso, ¿qué porcentaje de las pólizas tendrán que pagar una reclamación?

5. Un transistor tiene una distribución de tiempo de falla exponencial con tiempo medio de falla de 20 000 horas. El transistor ha durado 20 000 horas en una aplicación particular, ¿cuál es la probabilidad de que el transistor falle a las 30 000 horas?

6. Un transbordador lleva a sus clientes a través de un río cuando han abordado 10 automóviles. La experiencia muestra que que los autos arriban al transbordador de manera independiente a una tasa media de 7 por hora. Obtenga la probabilidad de que tiempo entre viajes consecutivos sea al menos de 1 hora.

7. Una caja de caramelos contiene 24 barras. El tiempo entre pedidos por barra se distribuye exponencialmente con media de 10 minutos, ¿cuál es la probabilidad de que una caja abierta a las 8:00 am se haya terminado al medio día?

8. El precio que se pide por cierto seguro de distribuye normalmente con media de \$50.00 y desviación estándar de \$5.00 Los compradores están dispuestos a pagar una cantidad también que se distribuye normalmente con media de \$45.00 y desviación estándar de \$2.50, ¿cuál es la probabilidad de que la transacción se realice?

9. La dureza de una aleación particular de distribuye normalmente con media de 70 y desviación estándar de 4. Si un espécimen se acepta sólo si su dureza está en el intervalo $(70 - c, 70 + c)$, ¿para qué valor de c el 95 % de los especímenes tendrían una dureza aceptable?

10. Un ciento de pequeños tornillos se empacan en una caja. Cada tornillo pesa 1 onza con desviación estándar de 0.01 onzas. Encuentre la probabilidad de que la caja pese más de 102 onzas.

11. El mecanismo interno de un refrigerador de cierto tipo tiene una vida cuya distribución es aproximadamente normal, con media $\mu = 12$ años y con desviación estándar de $\sigma = 4.863$ años. El fabricante asume la responsabilidad de reponer (o, en su caso,

reparar gratuitamente) aquellos refrigeradores que estén dentro de la garantía. Si piensa reponer solo el 5 % de las unidades, ¿por cuánto tiempo debe estipular la garantía?

12. Muchos hoteles a menudo aseguran reservaciones por encima de su capacidad, con objeto de minimizar las pérdidas ocacionadas por los clientes que no se presentan. Suponga que los archivos de un hotel (EL PASO) indican que, en promedio, 10 % de sus clientes no se presentan a reclamar sus reservaciones. Si el hotel acepta 215 reservaciones y solo hay 200 habitaciones en el hotel, ¿cuál es la probabilidad de que todos los clientes que lleguen a reclamar su reservación consigan una habitación?

13. Si 30 % de todos los estudiantes de ingeniería que ingresan a la universidad desertan durabtes el primer año, ¿cuál es la probabilidad de que mású de 600 de los 1 800 estudiantes de nuevo ingreso, para este año, deserten durante el primer año?

14. Las licenciaturas en ciencias puras en México (matemáticas, física, astronomía y biología) han captado cada vez menos estudiantes, y la enorme mayoría de universidades e institutos tecnológicos definitivamente no incluyen estas profesiones en su matrícula. Un estudio publicado en la revista del Consejo Nacional de Ciencia y Tecnología (octubre de 1 999) reveló qu 77 % de los estudiantes que inician una carrera en ciencias puras desertan durante los primeros dos años, y sólo 23 % se gradúan en un lapso no mayor a seis años. Si en 2 001 se registraron en todo el país 199 estudianes recién inscritos en alguna carrera de ciencias puras, determine:

    a) El número más probable de estudantes que se graduarán antes de que transcurran seis años.

    b) La probabilidad de que más de las tres cuartas partes de ellos deserten durante los primeros dos años.

    c) Si sólo 18 % de la matrícula en ciencias puras es de matemáticas, determine (de ese conjunto de 199) cuál es el número más probable de nuevos matemáticos que se graduarán en México durante los próximos seis años.

Capítulo 6

Estadística descriptiva

En esta sección se estudia los conceptos básicos de la estadística, así como la organización, resumen y presentación de datos que en un primer curso de estadística se debe estudiar, a esta primera parte se le conoce como estadística descriptiva. (ver [5]).

Definición 6.0.1.

1. *Una Medición es un número o categoría derivado de una observación de un fenómeno, con el objetivo de expresarlo.*

2. *Un Elemento de medición es un objeto en el cual se toman las mediciones.*

3. *Una Población es una colección de elementos de medición.*

4. *Una Muestra es un subconjunto de una población.*

5. *Un Dato es una medición.*

Definición 6.0.2. *Estadística. Arte y Ciencia de reunir, analizar, presentar e interpretar datos, y se divide en: Estadística descriptiva y en Estadística inferencial.*

$$\text{Estadística}\begin{cases}\text{Estadística descriptiva}\\ \text{Estadística inferencial}\end{cases}$$

Definición 6.0.3. *Estadística descriptiva. Rama de la Estadística que trata de la organización, resumen y presentación de datos.*

6.1. Clasificación de datos

Los datos para su estudio se clasifican en:

1. Cuantitativos. Son aquellos datos numéricos con propiedades algebraicas de interés.

2. Ordinales. Son aquellos datos, donde su propiedad de orden proporciona clases o jerarquias.

3. Nominales. Son aquellos que no son Cualitativos, ni Ordinales.

$$\text{Clasificación de datos}\begin{cases}\text{Cualitativos}\begin{cases}\text{Nominales}\\ \text{Ordinales}\end{cases}\\ \text{Cuantitativos}\end{cases}$$

Ejemplo 6.1.1.

Dato	*Clasificación*
Nombre	*Nominal*
Edad	*Cuantitativo*
Género	*Nominal*
Número telefónico	*Nominal*
Peso	*Cuantitativo*
Estado civil	*Nominal*
Escolaridad	*Ordinal*

6.2. Representación de datos

$$\text{Datos}\begin{cases}\text{Tabla de frecuencias}\\ \text{Gráficos}\begin{cases}\text{Pastel(Cualitativos)}\\ \text{Barras(Cualitativos)}\\ \text{Histograma(Cuantitativos)}\\ \text{Polígono de frecuencias(Cuantitativos)}\\ \text{Ojiba(Cuantitativos)}\end{cases}\end{cases}$$

6.3. Medidas de tendencia central

Definición 6.3.1. *Las Medidas de tendencia central son características de una población. Las más usadas son:*

1. *Media*

   a) *La Media aritmética $\bar{x}$ de la muestra $\{x_1, x_2, \ldots, x_n\}$, se define como:*

   $$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

   b) *La Media ponderada $\bar{x}_P$ de la muestra $\{x_1, x_2, \ldots, x_n\}$ con pesos respectivos $p_1, p_2, \ldots, p_n$, se define como:*

   $$\bar{x}_P = \frac{x_1p_1 + x_2p_2 + \cdots + x_np_n}{p_1 + p_2 + \cdots + p_n}$$

2. *La Mediana M_d de la muestra $\{x_1, x_2, \ldots, x_n\}$, se define como el valor central de los datos ordenados. Así, si suponemos que $y_1 \leq y_2 \leq \cdots \leq y_n$ como ordenación de los datos de la muestra, entonces la mediana de dicha muestra será,*

$$M_d = \begin{cases} y_{\frac{n+1}{2}} & \text{si } n \text{ es impar,} \\ \\ \dfrac{y_k + y_{k+1}}{2} & \text{si } n \text{ es par, done } k = n/2. \end{cases}$$

En otras palabras la Mediana M_d es el valor de la población donde se acumula el 50 % de los datos ordenados.
En el caso de tener datos no Cuantitativos y trabajar con una muestra de tamaño impar, elimine o genere un dato al azar para poder utilizar la fórmula anterior (caso n impar).

3. *La Moda M_o de una muestra, se define como el dato con mayor frecuencia.*

Nota 6.3.2.

La Media aritmética de una población μ, es una característica numérica de poblaciones de datos cuantitativos.

La Mediana de una población, es una característica de poblaciones de datos Ordinales.

La Moda de una población es una característica de poblaciones de datos Nominales.

Ejemplo 6.3.3. *A continuación se muestran datos de contenido de cafeína en tabletas de diferentes laboratorios:*

Laboratorio	*mg de cafeína por tableta*	*# de tabletas*
Lab. 1	*150*	*5*
Lab. 2	*205*	*3*
Lab. 3	*180*	*1*
Lab. 4	*250*	*2*
Lab. 5	*300*	*4*

Si se mezclan las 15 tabletas, ¿cuál es el contenido medio de cafeína por tableta?

$$\bar{x}_P = \frac{(150)(5) + (205)(3) + (180)(1) + (250)(2) + (300)(4)}{5+3+1+2+4} = \frac{3245}{15} = 216.\bar{3}$$
luego el contenido medio de cafeína por tableta es de $216.\bar{3}$ *mg*.

Ejemplo 6.3.4. *Se quiere probar la toxicidad de un insecticida, para ello se toma una gran cantidad de insectos y se les rocía de insecticida, observando que aproximadamente 20 % de ellos mueren inmediatamente; 45 minutos después mueren 30 % más y 3 horas más tarde mueren 15 %; el resto no muere. ¿Cuál es la medida de tendencia central adecuada para saber la vida media de estos insectos cuando son rociados con el insecticida? ¿Cuál es el valor de esta medida?*

Estos datos son tiempos de vida los cuales son datos cuantitativos, pero no tenemos información suficiente (por ejemplo del 35 % que no muere) para hallar la media aritmética ponderada. Estos datos como son cuantitativos también son datos ordinales y la medida de tendencia central adecuada para estos datos es la mediana,
$M_d = 45$ minutos ya que en este valor se acumula el 50 % de los datos ordenados.

6.4. Medidas de dispersión

Definición 6.4.1.

1. *El Rango R de la muestra $\{x_1, x_2, \ldots, x_n\}$, es el menor de los intervalo donde se encuentra la muestra, por lo que:*
$$R = [m, M]$$
donde $m = \text{mín}\{x_1, x_2, \ldots, x_n\}, \quad M = \text{máx}\{x_1, x_2, \ldots, x_n\}$.

2. *La Varianza S^2 de la muestra $\{x_1, x_2, \ldots, x_n\}$, se define como:*
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

3. *La Desviación estandar S de la muestra $\{x_1, x_2, \ldots, x_n\}$, se define como la raiz cuadrada de su Varianza.*
$$S = \sqrt{Varianza}$$

Nota 6.4.2. *La Varianza nos mide la dispersión de los datos con respecto de la Media.*
$$\text{(Varianza cero)} \;\; S^2 = 0 \quad \Longleftrightarrow \quad x_i = \bar{x}, \; i = 1, \ldots x_n \;\; \text{(No hay dispersión).}$$

6.5. Gráficas de Barra y de Pastel

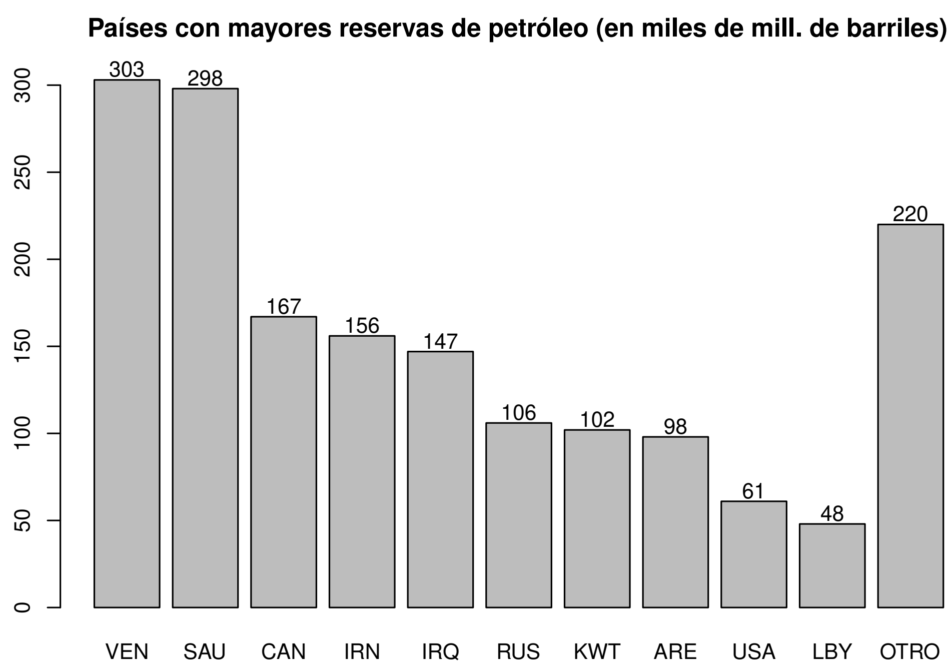
Las gráficas idóneas para datos Nominales y Ordinales son la de barra y la de Pastel.

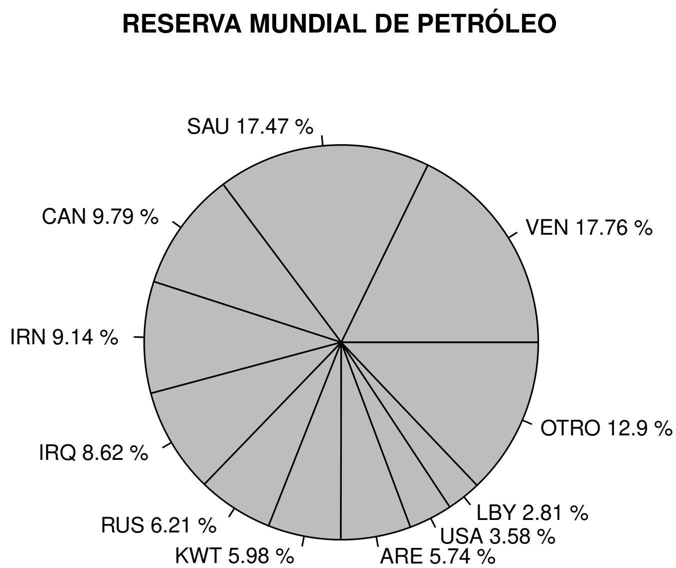
Ejemplo 6.5.1. *Reserva mundial de petróleo.*

Países con mayores reservas de petróleo en 2018 (en miles de mill. de barriles)	
Venezuela (VEN)	*303*
Arabia Saudita (SAU)	*298*
Canadá (CAN)	*167*
Irán (IRN)	*156*
Irak (IRQ)	*147*
Rusia (RUS)	*106*
Kuwait (KWT)	*102*
Emiratos Árabes Unidos (ARE)	*98*
Estados Unidos (USA)	*61*
Libia (LBY)	*48*
Otros (OTRO)	*220*

Instrucciones en *R* que generan las siguientes gráficas de Barras y de Pastel para el Ejemplo 6.5.1.

```
>            #--DATOS
> Reserva <- c(303,298,167,156,147,106,102,98,61,48,220)
> Pais<-c("VEN","SAU","CAN","IRN","IRQ","RUS","KWT","ARE","USA","LBY","OTRO")
>            #--GRAFICA DE BARRAS
> barras<-barplot(Reserva,names.arg=Pais,col=rainbow(length(Pais)),main="Paises
con mayores reservas de petroleo (en miles de mill. de barriles)")
> text(barras,Reserva+6,labels=Reserva, xpd=TRUE)
>             #--GRAFICA DE PASTEL
> frecrelativax100 <- round(Reserva/sum(Reserva)*100,2)
> ETIQUETA <- paste(Pais,frecrelativax100,"%","")
> pie(Reserva,labels=ETIQUETA,col=rainbow(length(Pais)),
  main="RESERVA MUNDIAL DE PETROLEO")
```





6.6. Datos agrupados

La siguente tabla nos indica que de una muestra de 80 elementos, $f_1 = 8$ de ellos pertenecen a la primera clase (intervalo) $140 < x \leq 150$,...., y , $f_6 = 17$ de ellos pertenecen a la última clase o intervalo $190 < x \leq 200$; las marcas de clase m_i son los puntos intermedios de cada clase; y las F_i indican las frecuencias acumuladas en cada clase.

TABLA DE FRECUENCIAS

Clase	marca de clase (m_i)	frecuencia (f_i)	frecuencia acumulada (F_i)
$140 < x \leq 150$	145	8	8
$150 < x \leq 160$	155	12	20
$160 < x \leq 170$	165	17	37
$170 < x \leq 180$	175	5	42
$180 < x \leq 190$	185	21	63
$190 < x \leq 200$	195	17	80

Partiendo de la información anterior conocida como tabla de frecuencias, un análisis de datos agrupados consiste en obtener las aproximaciones de la media aritmética, mediana, moda, varianza y desviación estándar de los 80 datos originales, utilizando como muestra las 80 marcas de clase: m_i con frecuencia f_i, $i = 1, 2, \ldots, 6$.

Media aritmética

$$\overline{\boldsymbol{X}} \approx \frac{\boldsymbol{m_1 \cdot f_1 + \cdots + m_K \cdot f_K}}{\boldsymbol{n}}$$

m_i: marcas de la clase i.
f_i: frecuencia de la clase i.
n: total de datos.

$$\overline{X} \approx \frac{145(8)+155(12)+165(17)+175(5)+185(21)+195(17)}{80} = 173.75$$

Mediana La aproximación viene dada por la mediana de $n = 80$ datos distribuidos de la siguiente manera: los f_i datos de la clase i son distribuidos uniformemente en dicha clase.

$$\boldsymbol{M_d \approx l_{M_d} + \left(\frac{\frac{n}{2} - F}{f}\right)(L)}$$

l_{M_d}: límite inferior de la clase donde se encuentra la Mediana.
F: frecuencia acumulada en la clase anterior donde se encuentra la Mediana.
f: frecuencia de la clase donde se encuentra la Mediana.
L: longitud de clase y n: es el total de datos.

Como la Mediana es el valor donde se acumula el 50 % de los datos ordenados, entonces la mediana es el dato ordenado $\frac{n}{2}$, luego, la clase i donde ella se encuentra es aquella que satisfface:

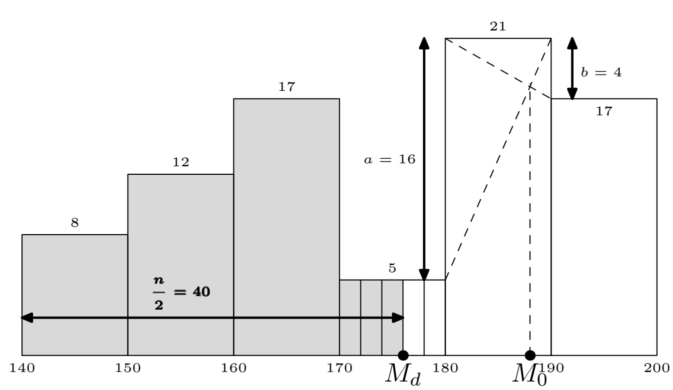
$$0 \leq F_{i-1} < n/2 \leq F_i$$

TABLA DE FRECUENCIAS

Clase	m_i	f_i	F_i
			$37 < n/2 \leq 42$
$160 < x \leq 170$	165	17	37
$170 < x \leq 180$	175	5	42

por lo que: $l_{M_d} = 170$, $f = 5$, $F = 37$, $L = 10$.

$$M_d \approx 170 + \left(\frac{40-37}{5}\right)(10) = 176$$



Moda La aproximación viene dada por la marca de clase con mayor frecuencia, pero daremos una mejor aproximación recargando esta marca de clase del lado con mayor frecuencia (ver figura).

Por semejanza de triángulos se tiene, $\dfrac{a}{b} = \dfrac{M_o - l_{M_o}}{L - M_o + l_{M_o}}$, por lo que:

$$\boldsymbol{M_o \approx l_{M_o} + \left(\tfrac{a}{a+b}\right)(L)}$$

l_{M_o}: límite inferior de la clase donde se encuentra la Moda.
a: frecuencia modal - la frecuencia anterior.
b: frecuencia modal - la frecuencia posterior.
L: longitud de clase.

TABLA DE FRECUENCIAS

Clase	m_i	f_i	F_i
$170 < x \leq 180$	175	5	42
$180 < x \leq 190$	185	21	63
$190 < x \leq 200$	195	17	80

Por lo que $l_{M_o} = 180$, $a = 21 - 5$, $b = 21 - 17$, $L = 10$, entonces,

$$\begin{aligned} M_o \approx& 180 + \left(\frac{16}{16+4}\right)(10) \\ =& 180 + \frac{16}{20}(10) = 180 + 8 = 188 \end{aligned}$$

Varianza

$$\begin{aligned} S^2 =& \frac{1}{n-1}\sum_{i=1}^{k}[(m_i - \bar{x})^2 \cdot f_i] = \frac{1}{79}[(145 - 173.75)^2(8) + (155 - 173.75)^2(12) \\ &+ (165 - 173.75)^2(17) + (175 - 173.75)^2(5) + (185 - 173.75)^2(21) \\ &+ (195 - 173.75)^2(17)] = 284.4937 \end{aligned}$$

Las muestras de datos cuantitativos se pueden representar por una tabla de frecuencias, y con tres gráficas: Histograma, Polígono de frecuenciad y Ojiva.
En el siguiente ejemplo realizaremos un análisis estadístico descriptivo completo con el uso del paquete estadístico R.

Ejemplo 6.6.1. *La siguiente tabla nos muestra la resistencia a la tensión, en libras por pulgada cuadrada, de 80 especímenes de una nueva aleación de aluminio y litio, que está siendo* *evaluada como posible material para su uso en la fabricación de aeronaves.*

105	221	183	186	121	181	180	143
97	154	153	174	120	168	167	141
245	228	174	199	181	158	176	110
163	131	154	115	160	208	158	133
207	180	190	193	194	133	156	123
134	178	76	167	184	135	229	146
218	157	101	171	165	172	158	169
199	151	142	163	145	171	148	158
160	175	149	87	160	237	150	135
196	201	200	176	150	170	118	149

I. **Tabla de frecuencias** Las siguientes tres reglas se utilizan frecuentemente para hallar K (número de clases en una tabla de frecuencias),

   a) $5 \leq K \leq 20$.

   b) $K \approx \sqrt{n}$.

   c) $K \approx 1 + \log_2(n)$ (Regla de Sturges).

   Instrucción en R, para hallar el rango de la muestra "Datos".

```
> Datos <- c(105,221,183,186,121,181,180,.......,170,118,149)
> range(Datos)
[1]  76 245
```

   Como $\sqrt{80} \approx 8.9443$, $K \approx 1 + \log_2(80) \approx 7.3219$, y Rango* $R = [76, 245]$.

   Para hallar la longitud de las clases con las que vamos a trabajar dividamos el Rango en K subintervalos

K	(longitud de R)/K	L
6	$169/6 = 28.1666$	29
7	$169/7 = 24.1429$	25
8	$169/8 = 21.125$	22
9	$169/9 = 18.77$	19
10	$169/10 = 16.9$	17

   Analizando la tabla anterior y considerando que lo que se quiere es la mejor representación gráfica, observamos que podemos trabajar con: 6 clases de longitud 30, ó 7 clases de longitud 25, ó 9 clases de longitud 20.

   Trabajaremos con $K = 9$ clases, de longitud $L = 20$, iniciando en $I = 70$.

Instrucción en R, para generar la Tabla de frecuencias.

```
> h <- hist(Datos,breaks=70+20*(0:9),right = TRUE,plot = FALSE)
> Tabla.de.frecuencias <- data.frame(Clases=cut(h$mids,breaks=70+20*(0:9)),
  mi=h$mids,fi=h$counts,Fi=cumsum(h$counts))
> Tabla.de.frecuencias
     Clases  mi fi Fi
1   (70,90]  80  2  2
2  (90,110] 100  4  6
3 (110,130] 120  5 11
4 (130,150] 140 16 27
5 (150,170] 160 21 48
6 (170,190] 180 17 65
7 (190,210] 200  9 74
8 (210,230] 220  4 78
9 (230,250] 240  2 80
```

Análisis de datos agrupados

i) Media aritmética

$$\begin{aligned}\overline{X} \approx& \frac{m_1 \cdot f_1 + \cdots + m_K \cdot f_K}{n}\\ =& \frac{(80)(2) + (100)(4) + (120)(5) + \cdots + (220)(4) + (240)(2)}{80}\\ =& \frac{160 + 400 + 600 + 2240 + 3360 + 3060 + 1800 + 880 + 480}{80}\\ =& \frac{12980}{80}\\ =& \boxed{162.25 \ \ \text{lb/plg}^2}\end{aligned}$$

ii) Mediana

$$\begin{aligned}M_d \approx& l_{M_d} + \left(\frac{\frac{n}{2} - F}{f_i}\right)(L)\\ =& 150 + \left(\frac{40 - 27}{21}\right)(20)\\ =& 150 + \left(\frac{13}{21}\right)(20)\\ =& 150 + \frac{260}{21} = \frac{3410}{21} = \boxed{162.38095 \ \ \text{lb/plg}^2}\end{aligned}$$

iii) Moda

$$\begin{aligned}
M_o \approx & l_{M_o} + \left(\frac{a}{a+b}\right)(L) \\
= & 150 + \left(\frac{5}{5+4}\right)(20) \\
= & 150 + \left(\frac{5}{9}\right)(20) \\
= & 150 + \frac{100}{9} = \frac{1450}{9} = \boxed{161\ .\overline{1}\ \text{ lb/plg}^2}
\end{aligned}$$

MEDIDAS DE DISPERSIÓN

i) Varianza

$$S^2 = \text{Varianza} \approx \frac{\sum_{i=1}^{K}(m_i)^2(f_i) - (n)(\overline{X})^2}{n-1}$$

$$= \frac{(m_1)^2 \cdot f_1 + \cdots + (m_K)^2 \cdot f_K - (n)(\overline{X})^2}{n-1}$$

$$= \frac{(80)^2(2) + (100)^2(4) + (120)^2(5) + \cdots + (220)^2(4) + (240)^2(2)}{80}$$

$$= \frac{12800 + 40000 + 72000 + 313600 + 537600 + 550800 + 360000 + 193600 + 115200 - (80)(162.25)^2}{79}$$

$$= \frac{2195600 - 2106005}{79} = \frac{89595}{79}$$

$$= \boxed{1134.11392\ \ (\text{lb/plg}^2)^2}$$

ii) Desviación estandar

$$S = \text{Desviación estandar} = \sqrt{\text{Varianza}} \approx \sqrt{1134.11392}$$

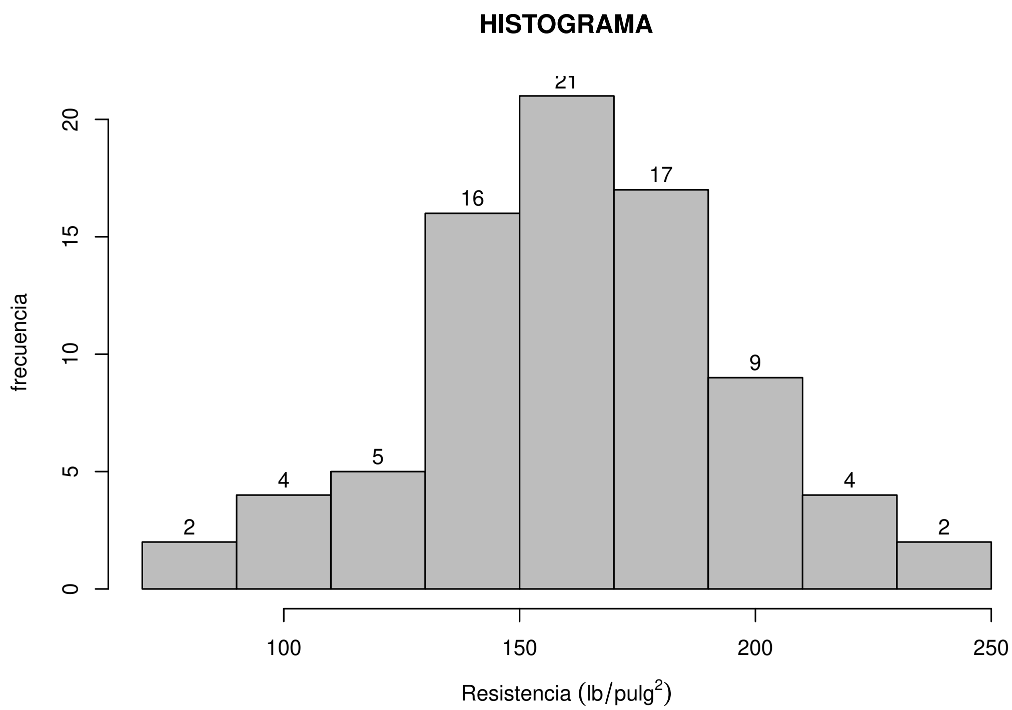
$$= \boxed{33.6766\ \ \text{lb/plg}^2}$$

El Histograma, la Ojiva y el Polígono de frecuencias, son las representaciones gráficas de una Tabla de frecuencias.

6.7. Histograma, Polígono de frecuencias y Ojiva

Instrucciones en R para obtener el correspondiente Histograma del Ejemplo 6.6.1.

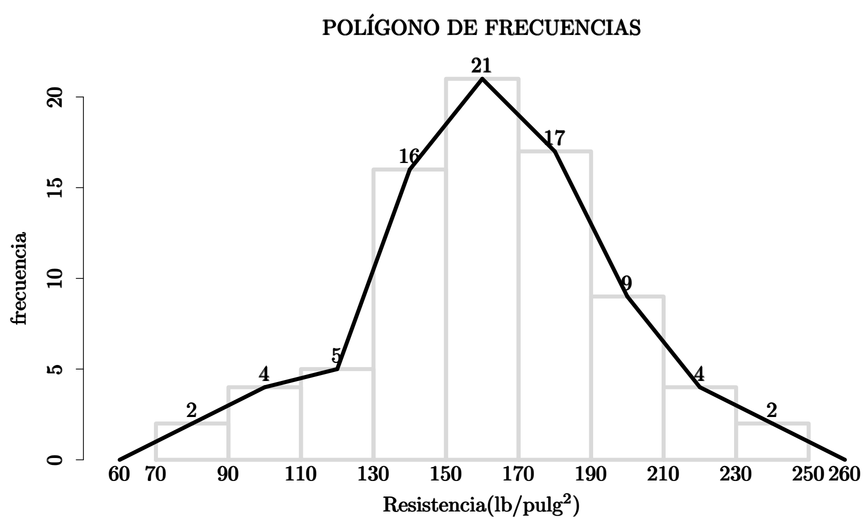
```
> Datos <- c(105,221,183,186,121,181,180,..............,170,118,149)
> hist(Datos,breaks=70+20*(0:9),right = TRUE,xlab = expression(paste
  ("Resistencia",(lb/pulg^2))),ylab = "frecuencia",main = "HISTOGRAMA",
  col=rainbow(9),labels = TRUE)
```



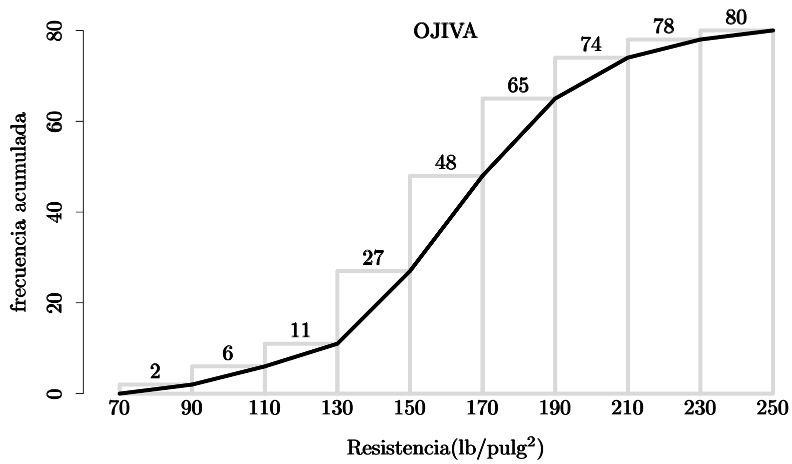
El Polígono de frecuencias tiene como vértices al conjunto de puntos

$$\{(m_1 - L, 0); (m_1, f_1); (m_2, f_2); \ldots; (m_k, f_K); (m_K + L, 0)\}$$

donde K es el total de clases y $m_i,\ i = 1, 2, \ldots, K$ son las marcas de clase.



La Ojiva es la curva no decreciente que se obtiene uniendo con segmentos de rectas los puntos $\{(L_1 - L, 0), (L_1, F_1), (L_2, F_2), \ldots, (L_K, F_K)\}$. Donde K es el total de clases, L longitud de las clases, y para $i = 1, 2, \ldots, K$, L_i y F_i son respectívamente los límites superiores y la frecuencia acumulada de la clases i.



6.8. Ejercicios

1. Para una muestra con mil elementos, completar la siguiente tabla y graficar, el histograma.

Clase		m	f_i	f_i^*	F_i	F_i^*
			4			
			26		38	
						0.091
				0.089		
			146			
154.5	157.5				630	
				0.125		
						0.88
					972	
			22			
						0.998
				0.001		
			1			

2. Los datos mostrados a continuación son los sueldos de inicio de una muestra aleatoria de 100 estudiantes de computación o de sistemas de cómputo que recibieron sus grados

de bachillerato durante 1993

Sueldos de inicio (en miles de pesos)									
24.2	29.9	23.4	23.0	25.5	22.0	33.9	20.4	26.6	24.0
28.9	22.5	18.7	32.6	26.1	26.2	26.7	20.4	22.2	24.7
18.6	18.5	19.6	24.4	24.8	27.8	27.6	27.2	20.8	22.1
19.7	25.3	28.2	34.2	32.5	30.8	26.8	20.6	21.2	20.7
25.2	25.7	32.2	28.8	24.7	18.7	20.5	25.5	19.1	25.5
22.1	27.5	25.8	25.2	25.6	25.2	25.2	27.9	18.9	37.3
29.9	23.2	19.8	20.8	29.5	27.6	21.2	38.7	21.3	24.8
32.3	20.1	26.8	25.4	26.3	21.2	19.5	22.8	21.7	25.3
32.3	28.1	27.5	25.3	19.3	27.4	26.4	20.9	34.5	25.9
31.4	27.4	27.3	20.6	31.8	25.8	25.2	21.9	26.8	26.5

3. En la siguiente tabla se muestran los datos a la resistencia en libras por pulgada cuadrada de 100 botellas de vidrio no retornables de un litro de refresco, los cuales se obuvieron probando cada botella hasta romperla.
   Elabore la tabla de frecuencias y utilícela para aproximar la media aritmética, la mediana, la moda, la varianza, la desviación estándar, construya el histograma.

Resistencia en $1b/in^2$
al rompimiento de 100 botellas de vidrio

265	197	346	280	265	200	221	265	261	278
205	286	317	242	254	235	176	262	248	250
263	274	242	260	281	246	248	271	260	265
307	243	258	321	294	328	263	245	274	270
220	231	276	228	223	296	231	301	337	298
268	267	300	250	260	276	334	280	250	257
260	281	208	299	308	264	280	274	278	210
234	265	187	258	235	269	265	253	254	280
299	214	264	267	283	235	272	287	274	269
215	318	271	293	277	290	283	258	275	251

Capítulo 7

Estadística inferencial

En esta sección empezamos el estudio de la estadística inferencial con la estimación puntual y por intervalo de confianza de la media aritmética, la cual es una característica importante de cualquier población numérica (ver [4], [5]).

Definición 7.0.1. *Estadística inferencial. Rama de la estadística que comprende métodos y procedimientos para deducir propiedades (hacer inferencias) de una población, a partir del análisis o esutudio de una muestra.*

$$\text{Estadística inferencial}\begin{cases}\text{Estimación}\begin{cases}\text{Puntual}\\ \text{Intervalo de confianza}\end{cases}\\ \text{Prueba de hipótesis}\end{cases}$$

7.1. Estimación paramétrica

Nota 7.1.1. *Sin pérdida de generalidad podemos suponer que la población esta formada por el conjunto de mediciones, la cual son valores que toma cierta variable aleatoria X, así, la distribución de datos observados en el polígono de frecuencias y en la ojiva, no son más que aproximaciones en cierta escala de las funciones de distribución f y F respectivamente, de esta variable aleatoria X. Por ésto, toda Población, la trabajaremos como variable aleatoria.*

Ejemplo 7.1.2. *La Población $\boldsymbol{X}$ de hombres (H) y mujeres (M) de la CDMX donde el $100p\,\%$ son mujeres, tiene distribución Bernoulli con parámetro p, ya que podemos observar al dato M como 1, y al dato H como 0 (M≡1 y H≡0).*

$$\textit{La población (de hombres y mujeres de la CDMX) } \boldsymbol{X} \sim B(p).$$

Nota 7.1.3. *Una muestra aleatoria de tamaño n de la población X, es un vector aleatorio $(X_1, X_2, \ldots, X_n)$, donde las variables aleatorias $X_i,\ i = 1, 2, \ldots, n$ son independientes y tienen la misma distribución que la población X.*

Definición 7.1.4. *Un parámetro θ de una población, es una característica numérica de la población.*

Ejemplo 7.1.5. *La Población X de hombres (H) y mujeres (M) de la CDMX, tiene distribución Bernoulli con parámetro p, donde p indica que el $100p\,\%$ de los habitantes de la CDMX son mujeres (M), así, 1≡M y 0≡H.*

$X \sim B(p)$, *la población de hombres tiene distribución Bernoulli, con parámetro p.*

Nota 7.1.6. *Una muestra aleatoria de tamaño n de la población X, es un vector aleatorio $(X_1, X_2, \ldots, X_n)$, donde las variables aleatorias $X_i,\ i = 1, 2, \ldots, n$ son independientes y tienen la misma distribución que la población X.*

Definición 7.1.7. *Un parámetro θ de una población, es una característica numérica de la población.*

Definición 7.1.8. *Un Estadístico ó Estimador$\hat{\theta}$ del parámetro θ de una población X,es una función h de una muestra aleatoria $(X_1, X_2, \ldots, X_n)$, es decir, una variable aleatoria de la forma,*

$$\hat{\theta} = h(X_1, X_2, \ldots, X_n)$$

7.1.1. Estimación puntual

Definición 7.1.9. *Una estimación puntual del parametro θ, de la variable aleatoria $X(\theta)$, es un valor $h(x_1, x_2, \ldots, x_n)$ de la variable aleatoria $\hat{\theta} = h(X_1, X_2, \ldots, X_n)$.*

7.1.2. Características de un buen estimador

1. **Sesgo**. La variable $\hat{\theta}$, es un estimador insesgado del parámetro θ si:

$$E\ (\hat{\theta}) = \theta.$$

2. Consistencia.

A continuación damos una lista de parámetros poblacionales, así como sus correspondientes estimadores óptimos.

Parámetro	θ	Estimador $\hat{\theta}$
Media	μ	$\bar{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$
Varianza	σ^2	$S^2 = \dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$
Proporción	p	$\hat{p} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$

7.1.3. Estimación por intervalo de confianza

Definición 7.1.10. *El intervalo del* $100(1-\alpha)\,\%$ *de confianza del parámetro* θ*, se define como el intervalo aleatorio,*

$$[I, S]$$

donde el límte inferior de confianza I *y el límte superior de confianza* S *son estadísticos* $h(X_1, X_2, \cdots, X_n)$*, tales que,* $P[I \leq \theta \leq S] = 1-\alpha$*, es decir, el* $100(1-\alpha)\,\%$ *de la veces:* $i \leq \theta \leq s$*, para valores específicos* i, s *de los estadísticos* I, S *respectivamente, por lo cual, definimos como una estimación por intervalo de confianza del* $100(1-\alpha)\,\%$ *para el parámetro poblacional* θ*, al intervalo,*

$$[i, s]$$

α *y* $(1-\alpha)$ *son llamados coeficiente de confianza y nivel de confianza respectivamente.*

Teorema 7.1.11. *Sea* $(X_1, X_2, \ldots, X_n)$ *una muestra aleatoria de tamaño* n *de la población* X*, entonces, el intervalo del* $100(1-\alpha)\,\%$ *de confianza de la media poblacional* μ*, viene dado por:*

1. $[\bar{X} - Z_{\alpha/2}\ \sigma/\sqrt{n}\ ,\ \bar{X} + Z_{\alpha/2}\ \sigma/\sqrt{n}]$
   cuando la población $X \sim N(\mu, \sigma^2)$ *ó si la muestra es grande, es decir* $n \geq 30$.

2. $[\bar{X} - t_{\alpha/2,n-1}\ S/\sqrt{n}\ ,\ \bar{X} + t_{\alpha/2,n-1}\ S/\sqrt{n}]$
   cuando la población $X \sim N(\mu, \sigma^2)$ *con varianza* σ^2 *desconocida.*

Demostración.

Población X	Estadístico	Distribución
Muestra grande $(n \geq 30)$ ó $X \sim N(\mu, \sigma^2)$	$Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$Z \sim N(\mu, \sigma^2/n)$
$X \sim N(\mu, \ \sigma^2)$, σ^2 desconocida	$T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$T \sim t_{n-1}$

Por el Teorema del Límite Central, ó por la propiedad Reproductiva de la distribución normal (suma de variables normales independientes es una variable normal), el estadístico Z tiene distribución $N(\mu, \sigma^2/n)$. Por la propiedad reproductiva se puede probar que el estadístico S tiene distribución ji-cuadrada, y el estadístico T distribución t student con $n-1$ grados de libertad.

En cada caso, debido a que se conoce la distribución del estadístico Z, ó T, podemos hallar el intervalo de confianza para la media poblacional μ.

PRIMER CASO: Población X normal ó muestra grande $n \geq 30$, con varianza σ^2 conocida.

El punto porcentual Z_α se define como el valor de la variable Z, tal que, $P[Z \geq Z_\alpha] = \alpha$

Por lo que, $\quad 1-\alpha = P[-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}] = P\left[-Z_{\alpha/2} \leq \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}\right]$,

despejando μ, $\quad P[\bar{X} - Z_{\alpha/2}\ \sigma/\sqrt{n} \leq \mu \leq \bar{X} + Z_{\alpha/2}\ \sigma/\sqrt{n}] = 1-\alpha$.

Obteniendo el intervalo del $100(1-\alpha)\,\%$ de confianza para la media poblacional μ

$$\bar{X} - Z_{\alpha/2}\ \sigma/\sqrt{n} \leq \mu \leq \bar{X} + Z_{\alpha/2}\ \sigma/\sqrt{n}$$

SEGUNDO CASO: Población X normal, con varianza σ^2 desconocida.

El punto porcentual $t_{\alpha,n-1}$ se define como el valor de T, tal que, $\quad P[T \geq t_{\alpha,n-1}] = \alpha$

Por lo que, $\quad 1-\alpha = P[-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}] = P\left[-t_{\alpha/2,n-1} \leq \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \leq t_{\alpha/2,n-1}\right]$,

despejando μ, $\quad P[\bar{X} - t_{\alpha/2,n-1}\ S/\sqrt{n} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1}\ S/\sqrt{n}] = 1-\alpha$.

Obteniendo el intervalo del $100(1-\alpha)\,\%$ de confianza para la media poblacional μ

$$\bar{X} - t_{\alpha/2,n-1}\ S/\sqrt{n} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1}\ S/\sqrt{n}$$

□

Nota 7.1.12. *Por Teorema anterior el error de estimación $|\bar{X}-\mu|$ por intervalo de confianza de la media poblacional μ, satisface* $\quad |\bar{X}-\mu| \leq Z_{\alpha/2}\ \sigma/\sqrt{n}$, *por lo que,*

1. *el error máximo de estimación E, por intervalo del $100(1-\alpha)\,\%$ de confianza para la media poblacional μ, viene dado por,*

$$E = Z_{\alpha/2}\ \sigma/\sqrt{n}$$

2. *despejando, obtenemos el tamaño de muestra n, para una estimación por intervalo del $100(1-\alpha)\,\%$ de confianza para la media poblacional μ, con un error máximo E,*

$$n = \left(\frac{Z_{\alpha/2}\ \sigma}{E}\right)^2$$

Ejemplo 7.1.13. *Un ingeniero civil está probando la resistencia compresiva de concreto. Realiza la prueba con 16 especímenes, obteniendo los siguientes resultados. Suponga que la resistencia compresiva se distribuye aproximadamente normal.*
Obtenga una estimación por intervalo del 98 % de confianza para la media de resistencia compresiva.

Resistencia compresiva (lb/plg^2)			
2216	*2237*	*2249*	*2204*
2225	*2301*	*2281*	*2263*
2318	*2255*	*2275*	*2295*
2250	*2238*	*2300*	*2217*

Datos

$n = 16$
$\bar{x} = 2257.75$
$s = 34.51473$
$\alpha = 0.02$
$t_{\alpha/2,n-1} = 2.602$

Intervalo de confianza

$$\bar{X} - t_{\alpha/2,n-1}\ S/\sqrt{n} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1}\ S/\sqrt{n}$$

Estimación por intervalo

$$2257.75 - (2.602)(34.5147/4) \leq \mu \leq 2257.75 + (2.602)(34.5147/4)$$

$$\boxed{2235.29816\ lb/plg^2 \ \leq \mu \leq 2280.20183\ lb/plg^2}$$

Mostremos como se obtiene esta estimación, con el paquete estadístico *R*.

1. Primer método.

```
> RC <- c(2216,2237,2249,2204,2225,2301,2281,2263,...,2300,2217)
> t.test(RC,conf.level = 0.98)

 One Sample t-test

data:  RC
t = 261.66, df = 15, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
98 percent confidence interval:
 2235.294 2280.206
sample estimates:
mean of x
  2257.75
```

2. Segundo método: $\bar{x} = \text{mean(RC)}$, $s = \text{sd(RC)}$, $t_{\alpha/2,n-1} = \text{qt(0.99,15)}$

```
> RC <- c(2216,2237,2249,2204,2225,2301,2281,2263,...,2300,2217)
> i <- mean(RC)-qt(0.99,15)*sd(RC)/sqrt(16)
> s <- mean(RC)+qt(0.99,15)*sd(RC)/sqrt(16)
> intervalo.confianza <- c(i,s)
> intervalo.confianza
[1] 2235.294 2280.206
```

Ejemplo 7.1.14. *Se sabe que la vida en horas de una bombilla eléctrica de 100 watts se distribuye aproximadamente normal, con desviación estándar $\sigma = 25$ horas. Una muestra aleatoria de 36 bombillas tiene una vida media $\bar{x} = 1014$ horas.*

1. *Construya un intervalo de confianza del 95 % respecto a la vida media de las bombillas eléctricas.*
2. *Si queremos que en esta estimación, el error máximo sea de siete horas, ¿qué tamaño de muestra debe usarse?*

1. Datos: $n = 36$, $\bar{x} = 1014$ horas, $\sigma = 25$ horas, $\alpha = 0.05$, $Z_{\alpha/2} = 1.96$

Intervalo de confianza

$$\bar{X} - Z_{\alpha/2}\ \sigma/\sqrt{n} \leq \mu \leq \bar{X} + Z_{\alpha/2}\ \sigma/\sqrt{n}$$

Estimación por intervalo

$$1014 - (1.96)25/6 \leq \mu \leq 1014 + (1.96)\ 25/6$$

$$\boxed{1005.8\bar{3} \text{ hrs } \leq \mu \leq 1022.1\bar{6} \text{ hrs}}$$

Con el paquete *R*: $Z_{\alpha/2} = Z_{.025} = \text{qnorm}(0.975)$

```
> i <- 1014-qnorm(0.975)*25/sqrt(36)
> s <- 1014+qnorm(0.975)*25/sqrt(36)
> intervalo.confianza <- c(i,s)
> intervalo.confianza
[1] 1005.833 1022.167
```

2. Datos: $E = 7$ horas

Tamaño de muestra

$$n = \left(\frac{Z_{\alpha/2}\ \sigma}{E}\right)^2 = \left(\frac{1.96(25)}{7}\right)^2$$

$$\boxed{n = 49}$$

7.2. Pruebas de hipótesis paramétricas

Definición 7.2.1.

1. *Hipótesis paramétrica. Enunciado acerca de una característica numérica (parámetro) de la distribución de probabilidad de una variable aleatoria ó de una población.*
2. *Prueba de hipótesis paramétrica. Procedimiento de la toma de decisión de, aceptar ó rechazar una hipótesis paramétrica.*

Nota 7.2.2. *La prueba de hipóteis consiste en suponer que la población X, satisface la hipótesis nula $H_0 : \theta = \theta_0$, a continuación se hace un estudio de la población con ayuda de la información contenida en una muestra aleatoria $(X_1, X_2, \ldots, X_n)$, si en el análisis se dan inconsistencias o contradicciones con la hipótesis establecida $H_0 : \theta = \theta_0$, entonces, se dará la conclusión fuerte de rechazar H_0 ó equivalentemente aceptar $H_a : \theta \neq \theta_0$, en caso contrario se daría la conclusión débil de aceptar $H_0 : \theta = \theta_0$ ó equivalentemente: no hay suficiente evidencia de aceptar $H_a : \theta \neq \theta_0$.*

Esta inconsistencia o contradicción se mostrará cuando el valor del estadístico de prueba $h(x_1, x_2, \ldots, x_n) \in R$, *donde* R *es conocida como Región crítica o Región de rechazo de* H_0, *el estudio de la Región crítica se inicia en el Lema de Neyman-Pearson.*
Cabe mencionar que en toda prueba de hipótesis estadística existe la posibilidad de cometer dos tipos de errores, como se puede observar en la siguiente tabla.

	H_0 *verdadera*	H_0 *falsa*
Se acepta H_0	*No hay error*	*Error tipo II*
Se rechaza H_0	*Error tipo I*	*No hay error*

α, *y* β *denotan las probabilidades de cometer los errores de tipo I y de tipo II respectivamente, es decir,*

$$\alpha = P(\textit{Cometer error tipo I}\ ) = P(\textit{Rechazar } H_0 \mid H_0 \textit{ es verdadera})$$
$$\beta = P(\textit{Cometer error tipo II}) = P(\textit{Aceptar } H_0 \mid H_0 \textit{ es falsa})$$

α *se le conoce como nivel o significancia de la prueba y la fija el investigador al construir la región crítica, razón por la cual la conclusión de rechazar* H_0 *es una conclusión fuerte, ya que el único error que se podría cometer es el tipo I el cual es fijado por el investigador. La conclusión de aceptar* H_0 *es una conclusión débil ya que se podría cometer el error tipo II, esta conclusión se hará fuerte cuando calculemos el valor de* β *y éste sea pequeño.*

Resumen de pruebas de hipótesis para μ y σ^2, con nivel de significancia α			
Hipótesis nula	Est. de prueba	Hipótesis alternativa	Criterio de rechazo
$H_0 : \mu = \mu_0$. σ^2 conocida, X Normal ó ($n \geq 30$)	$Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$H_a : \mu < \mu_0$	$Z_0 < -Z_\alpha$
		$H_a : \mu \neq \mu_0$	$\|Z_0\| > Z_{\alpha/2}$
		$H_a : \mu > \mu_0$	$Z_0 > Z_\alpha$
$H_0 : \mu = \mu_0$. X Normal, σ^2 desconocida	$T_0 = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$H_a : \mu < \mu_0$	$T_0 < -t_{\alpha,n-1}$
		$H_a : \mu \neq \mu_0$	$\|T_0\| > t_{\alpha/2,n-1}$
		$H_a : \mu > \mu_0$	$T_0 > t_{\alpha,n-1}$
$H_0 : \sigma^2 = \sigma_0^2$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$H_a : \sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha,n-1}$
		$H_a : \sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha/2,n-1}$ ó $\chi^2 > \chi^2_{\alpha/2,n-1}$
		$H_a : \sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{\alpha,n-1}$

Ejemplo 7.2.3. *Un proceso de fabricación de jabón de tocador debe producir un promedio de 120 barras por lote. No se desea tener cantidades menores ni mayores que el estándar. Una muestra de 10 lotes dio como resultado las siguientes cantidades de barras de jabón. Se supone que la población es normal.*

108 118 120 122 119 113 124 122 120 123

Con el nivel de significancia α*=0.05, pruebe si los resultados de esta muestra indican que el proceso de manufactura está trabajando en forma correcta.*

Datos	Hipótesis	Est. de Prueba.	Criterio de rechazo
$n = 10$	$H_0 : \mu = 120$	$T_0 = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$\lvert T_0 \rvert > t_{\alpha/2,n-1} = 2.262157$
$\bar{x} = 118.9$	$H_a : \mu \neq 120$		
$s = 4.931757$		Valor del E.P.	
$\alpha = 0.05$		$T_0 = \dfrac{118.9 - 120}{4.931757/\sqrt{10}} = -0.705$	
$t_{\alpha/2,n-1} = 2.262157$			

Conclusión:

Como el valor del estadístico de prueba $T_0 = -0.7053277$ no satisface el criterio rechazo, entonces, no hay evidencia suficiente que indique que el proceso de manufactura esté trabajando en forma incorrecta.

Nota 7.2.4. *Usando el paquete R y el criterio:*

$$\textit{Si } \alpha > \textit{p-value, entonces se rechaza } H_0$$

```
> Datos <- c(108,118,120,122,119,113,124,122,120,123)
> t.test(Datos,mu=120,conf.level = .95)

 One Sample t-test

data:  Datos
t = -0.70533, df = 9, p-value = 0.4985
alternative hypothesis: true mean is not equal to 120
95 percent confidence interval:
 115.372 122.428
sample estimates:
mean of x
    118.9
```

Como $0.05 = \alpha \ngtr$ p-value $= 0.4985$, entonces, no hay evidencia suficiente que indique que el proceso de manufactura esté trabajando en forma incorrecta.

Ejemplo 7.2.5. *El automóvil Ford Taurus aparece como con gran rendimiento de combustible en carretera, con un promedio de 30 millas por galón (Guía del comprador de automóvil nuevo 1995, de la revista Motor Trend). Un grupo defensor de los intereses del consumidor lleva a cabo pruebas de rendimiento que buscan obtener evidencia estadística que demuestre que los fabricantes sobrestiman los rendimientos de sus automóviles. En una muestra de 50 pruebas de rendimiento con el Ford Taurus , de obtiene una media de consumo en carretera de 29.5 millas por galón y una desviación estándar de 1.8 millas por galón. ¿A qué conclusión se debe llegar partiendo de los resultados de la muetra? Use un nivel de significancia de 0.01*

Datos	Hipótesis	Est. de Prueba.	Criterio de rechazo
$n = 50$	$H_0 : \mu = 30$	$Z_0 = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$	$Z_0 < Z_\alpha = -2.326$
$\bar{x} = 29.5$	$H_a : \mu < 30$		
$s = 1.8$		Valor del E.P.	
$\alpha = 0.01$		$Z_0 = \dfrac{29.5 - 30}{1.8/\sqrt{50}} = -1.964185$	
$Z_\alpha = 2.326$			

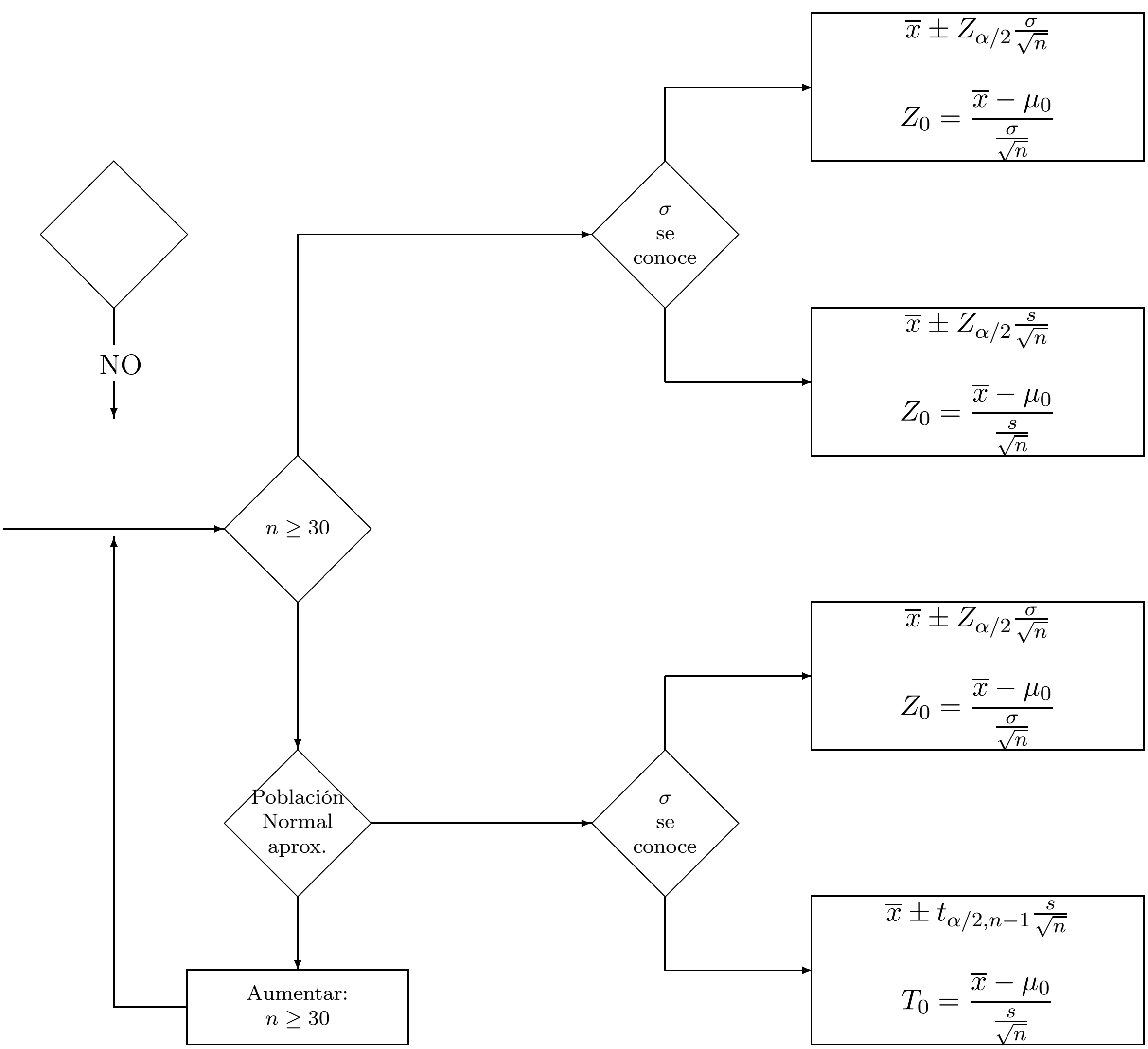
Conclusión:
Como el valor del estadístico de prueba $Z_0 = -1.964185$ no satisface el criterio de rechazo, entonces, no hay evidencia suficiente que indique que el fabricante sobrestime el rendimiento se sus automóviles.

Ejemplo 7.2.6. *Consideremos una máquina que se utiliza para llenar latas de refresco. Si la varianza del volumen de llenado excede* 0.02 u*, entonces, un gran porcentaje inaceptables de latas no se llenarán lo suficiente. Para monitorear la producción, el embotellador toma una muestra aleatoria de 20 latas, la cual produce una varianza de* $s^2 = 0.225\ u^2$*, ¿habrá suficiente evidencia de que la varianza de llenado exceda* 0.02 u*. Uitilice* $\alpha = 0.05$

Datos	Hipótesis	Est. de Prueba.	Criterio de rechazo
$n = 20$	$H_0 : \chi^2 = 0.02$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2 > 30.14$
$s^2 = 0.0225\ u^2$	$H_a : \chi^2 > 0.02$		
$\alpha = 0.05$		Valor del E.P.	
$\chi^2_{\alpha,n-1} = 30.14$		$\chi_0^2 = \dfrac{(19)(0.0225)}{0.02} = 21.375$	

Conclusión:
Como el valor del estadístico de prueba $\chi_0^2 = 21.375$ no cae en la región crítica, entonces, no hay evidencia suficiente de que la varianza de llenado excceda 0.02 u^2.

INTERVALOS DE CONFIANZA Y ESTADÍSTICOS DE PRUEBA PARA LA MEDIA POBLACIONAL μ



Estadístico de prueba	**Criterio de rechazo**		
	Cola inferior $H_a:\ \mu < \mu_0$	Dos colas $H_a:\ \mu \neq \mu_0$	Cola superior $H_a:\ \mu > \mu_0$
Z_0	$Z_0 < -Z_\alpha$	$\lvert Z_0\rvert > Z_{\alpha/2}$	$Z_0 > Z_\alpha$
T_0	$T_0 < -t_{\alpha,n-1}$	$\lvert T_0\rvert > t_{\alpha/2,n-1}$	$T_0 > t_{\alpha,n-1}$

Si el valor del estadístico de prueba satisface el Criterio de rechazo, entonces se rechaza la hipótesis H_o

7.3. Ejercicios

1. Para determinado modelo de automóvil se llevan a cabo pruebas de rendimiento de gasolina. Si la presición que se desea es un intervalo de confianza de 98 % con un margen de error de una milla por galón, ¿cuántos automóviles deben participar en la prueba? Suponga que las pruebas preliminares de rendimiento indican que la desviación estándar es de 2.6 millas por galón.

2. Al determinar la programación de las citas con pacientes, un centro médico desea una estimación de la media del tiempo que pasa un miembro de su personal con cada paciente. ¿De qué tamaño se debe tomar una muestra para que el margen de error deseado sea de 2 minutos con un nivel de confianza de 99 % Emplear un valor un valor de planeación de 8 minutos para la desviación estándar poblacional.

3. Una línea de producción funciona con una media del peso de llenado de 16 onzas por envase. El sobrellenado o la falta de llenado son problemas graves, y la línea de producción debe parar si se presenta alguno de ellos. De acuerdo a datos anteriores, se supone que el valor de σ es 8 onzas. Un inspector de calidad toma una muestra de 30 artículos cada dos horas, y de acuerdo con los resultados toma la decisión de parar la línea de producción para ajustes o dejarla trabajando. Si $\bar{x}$ =16.32 onzas, ¿qué acción recomendarías?

4. Una parte automotriz debe maquinarse de modo que se ajuste a las tolerancias que cumplen las exigencias de los clientes. Las especificaciones de producción requieren una varianza máxima en las longitudes de las partes de 0.0004. Suponga que la varianza de una muestra de 30 partes resulta ser $s^2 = 0.0005$. Con $\alpha = 0.05$, pruebe si se viola la especificación de la varianza poblacional.

5. Suponga que se va a implantar un nuevo método de producción si una prueba de hipótesis respalda la conclusión de que con este método se reduce la media del costo de operación por hora.

   Enuncie las hipótesis nula y alternativa si la media del costo para el método actual de producción es de 220 dólares por hora.

6. El voltaje de salida de cierto tipo de transformador de está investigando. Diez transformadores de este tipo se seleccionan al azar y se midió el voltaje dando como resultado $\overline{x}$=12.05 volts. Se sabe que la varianza de la población que es normal es de σ=0.8, construya un intervalo del 95 % de confianza para el voltaje medio.

7. Los agentes de ventas de autos TOYOTA tuvieron ventas anuales promedio de $1 000 000. Sebastián Durán, vicepresidente del negicio, propuso un plan de compensaciones con nuevos incentivos de venta. Cree que los resultados de un periodo de prueba le permitan llegar a la conclusión de que el plan de compensación aumente el promedio de ventas por agente.

   Elabore las hipótesis nula y alternativa adecuadas.

8. La Empresa de neumáticos Bridgeston afirma que sus neumáticos duran en promedio 28 000 millas. Las pruebas con 30 de sus neumáticos dan como resultados un promedio

de 27 500 millas de duración con una desviación estandar de 1 000 millas. Prueba la afirmación de la Empresa, con un nivel de significancia del .05

9. Un nuevo aditivo para la gasolina ha sido desarrollado por una compania norteamericana. En un experimento de uso del aditivo realizado en 8 autos durante un período de una semana se registraron los siguientes porcentajes de ahorros en el consumo de gasolina.

15.2	14.1	13.7	15.2
18.6	15.0	14.5	13.8

Use un intervalo de confianza del 95 % para estimar el ahorro promedio de gasolina.

10. Para estimar la media de consumo por cliente, en un gran restaurante, se reunieron datos de una muestra de 49 clientes. Suponga una desviación estándar poblacional de 5 dólares.

   a) Con nivel de confianza de 95 %, ¿cuál es el margen de error?

   b) Si la media de la muestra es de 24.80 dólares, ¿cuál es el intervalo de confianza de 95 %.

11. El alza en los precios de medicamentos recetados por los médicos provocó que le Congreso de Estados Unidos considerara leyes que obligaran a las compañias farmaceúticas a ofrecer descuentos a los ciudadanos carentes de beneficios en medicamentos. El Comité de reformas gubernamentales internas proporcionó datos acerca del costo de recetas para algunos fármacos de uso común. Suponga que los datos siguientes representan una muestra del costo de las recetas en dólares para un fármaco que se emplea para redicir el coloesterol.

110	112	115	99	100	98	104	126

Suponiendo que se trata de una población normal, ¿cuál es la estimación por intervalo de confianza del 95 % del costo promedio poblacional para este fármaco?

12. Suponga que se va a implantar un nuevo método de producción si una prueba de hipótesis respalda la conclusión de que con este método se reduce la media del costo de operación por hora.

   Enuncie las hipótesis nula y alternativa si la media del costo para el método actual de producción es de 220 dólares por hora.

13. El voltaje de salida de cierto tipo de transformador de está investigando. Diez transformadores de este tipo se seleccionan al azar y se midió el voltaje dando como resultado $\overline{x}$=12.05 volts. Se sabe que la varianza de la población que es normal es de $\sigma^2 = (0.8)^2$, construya un intervalo del 95 % de confianza para el voltaje medio.

14. Los agentes de ventas de autos TOYOTA tuvieron ventas anuales promedio de $1 000 000. Sebastián Durán, vicepresidente del negicio, propuso un plan de compensaciones con nuevos incentivos de venta. Sebastián Dirán que los resultados de un periodo de prueba le permitan llegar a la conclusión de que el plan de compensación aumente el promedio de ventas por agente.

a) Elabore las hipótesis nula y alternativa adecuadas.

b) ¿Cuál es el error de tipo I en este caso y cuál es la consecuencia de cometerlo?

c) ¿Cuál es el error de tipo II en este caso y cuál es la consecuencia de cometerlo?

15. Para determinado modelo de automóvil se llevan a cabo pruebas de rendimiento de gasolina. Si la presición que se desea es un intervalo de confianza de 98 % con un margen de error de 1 milla por galón, ¿cuántos automóviles deben participar en la prueba? Suponga que las pruebas preliminares de rendimiento indican que la desviación estándar es de 2.6 millas por galón.

Apéndice A

Tablas estadísticas

En esta sección construimos con C++ las tablas de las distribuciones *normal*, t, y $ji-cuadrada$.

A.1. Distribución normal estandar Z

A.1.1. Función: double pnorm(double x)

Esta función genera los valores de la función de distribución acumulada $\Phi(x)$ de la variable aleatoria normal estandar Z

$$\Phi(x) = P[Z \leq x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

En su obtención utilizamos el desarrollo en serie de Taylor de la función,

$$F(x) = \int_{0}^{x} e^{-y^2/2} dy = \sum_{n=0}^{\infty} \frac{x\ (-x^2/2)^n}{n!\ (2n+1)}$$

la cual es una serie: **casi** alternante.

```
double pnorm(double x)
{  int k;
   double  termino,suma,p,y;
        y=fabs(x);termino=y;suma=y;k=0;
           do{ k = k+1;  termino *=
               ((-y*y)*(2*k-1))/((2*k)*(2*k+1));  suma += termino;}
           while(fabs(termino > 0.00000001);
        if(x>=0) {p = 0.5+0.39894228 * suma;}
        else     {p = 0.5-0.39894228 * suma;}
   return p;
}
```

A.1.2. Función: double qnorm(double x)

Esta función genera los valores de la función $\Phi^{-1}(x)$, la cual se utiliza para generar la parte de la Tabla 2 de los valores de la distribución t con infinito grados de libertad.

```
double qnorm(double x)
{
  double q,a,b,e;
   a = -8.0;  b = 8.0;
   do{e = (b-a)/2;  q = (a+b)/2;
            if(pnorm(q) > x)
              {b = q;}
            else
               {  if(pnorm(q) < x)
         {a = q;}
         else{e = .00000001;};
                }
     }  while(e > 0.00000001);
    return q;
}
```

A.1.3. Tabla de la distribución Normal estándar

El siguiente programa genera la Tabla 1, correspondiente a valores de la función de distribución acumulada Φ, de la variable aleatoria normal estandar Z.

```
#include <stdio.h>
#include <math.h>
double pnorm(double x);
int main()
{  for (double i=-3.9;i<=0.01;i+=0.1)
            {for (double j=0.0;j<=0.09;j+=0.01)
                  {printf("& %lf \t",pnorm(i-j));} ;printf("\\\\ \n");}
   printf("CONTINUACION: X>=0 \n");
   for (double i=0.0;i<4.0;i+=0.1)
            {for (double j=0.0;j<=0.09;j+=0.01)
                 {printf("& %lf \t",pnorm(i+j));} ;printf("\\\\ \n");}
   return 0;
}
```

A.2. Distribución t

A.2.1. Función: double pt(double x,int k)

Esta función genera los valores de la función de distribución acumulada $F_t(x)$ de la variable aleatoria con distribución t con k grados de libertad.

$$F_t(x) = P[t \le x] = \int_{-\infty}^{x} f_k(y)\ dy = \frac{\Gamma[(k+1)/2]}{\sqrt{k\pi}\ \Gamma(k/2)} \int_{-\infty}^{x} \left(1+\frac{y^2}{k}\right)^{-(k+1)/2}$$

Como la función $f_k(x)$ es par, utilizando la sustitución trigonométrica $(x/\sqrt{k} = \tan\theta)$, obtenemos para $x \ge 0$

$$F_k(x) = \frac{1}{2} + \int_0^x f_k(y)\ dy = \frac{1}{2} + \frac{\Gamma[(k+1)/2]}{\sqrt{\pi}\ \Gamma(k/2)} \int_0^{\arctan(x/\sqrt{k})} \cos^{k-1}\theta\ d\theta$$

Para $n \ge 2$, escribiendo $\cos^n\theta = \cos^{n-2}\theta\ (1-\operatorname{sen}^2\theta)$ e integrando por partes el segundo término con $dv = \cos^{n-2}\theta \operatorname{sen}\theta\ d\theta$ y $u = \operatorname{sen}\theta$, obtenemos

$$\int_0^b \cos^n\theta\ d\theta = \frac{1}{n}\ \cos^{n-1}\theta\ \operatorname{sen}\theta\Big|_0^b + \frac{n-1}{n}\int_0^b \cos^{n-2}\theta\ d\theta$$

La siguiente función utiliza esta expresión recursiva para obtener los valores $F_k(x)$

```
double pt(double x, int k)  // x >= 0,  Ft(x,k) = P[T_k<=x] = p.
{ int n,j,d;
  double In,p,y,Gamas;
   y=fabs(x);
   d = k-(k/2)*2;
        if(d == 0)
           {j = 1; In = y/sqrt((y*y)+k);  Gamas = 0.5; }
        else
           {j = 0; In = atan(y/sqrt(k));  Gamas = 1.0/3.141593; }
   for(n = j+2; n<=k-1; n = n+2)
            { In = ((pow(sqrt(k),(n-1))*y)/(n*pow(sqrt((y*y)+k),n)))
                    +((((n-1)*In)/n)) ;
             Gamas = (n*Gamas)/(n-1);};
       if(x>=0) {p = 0.5+(Gamas* In);}
                      else     {p = 0.5-(Gamas* In);}
   return p;
 }
```

A.2.2. Función: double qt(double x,int k)

Esta función genera los valores de la función F_k^{-1} (Tabla 2), correspondiente a la distribución t con $k \ge 1$ grados de libertad.

```
double qt(double y, int k)  /* x >= 0,  Ft(x,k) = P[T_k<=x] = y, si y solo si  x = InvFt(x,k)
{ double a,b,e,x;
 if(k >= 3)
    {  a = -20.0;  b = 20.0;
     do{e = (b-a)/2;  x = (b+a)/2;
            if(pt(x,k) > y)  {b = x;}
            else
               {if(pt(x,k) < y)  {a = x;}
        else{e = 0.0000001;};
               }
       }  while(e > 0.0000001);
    }
     else
               {if(k == 2)
       {x=((2.0*y)-1) /sqrt((2.0*y)*(-y+1));}
       else{x=tan((3.141593/2)*((2.0*y)-1)) ;};
               }
  return x;
}
```

A.2.3. Tabla de la distribución t

El siguiente programa genera la Tabla 2.

```
double pnorm(double x);
double qnorm(double x);
double qt(double y, int k);
int main( )  /* TABLA 2 */
{ double v[]={.70,.75,.90,.95,.975,.99,.995,.9975,.999,.9995 };
    for (int k=1;k<=30;k++)
      {for (int i = 0; i<=9; i++)
     {printf("& %lf \t",qt(v[i],k));} ;printf("\\\\ \n");};
 for (int i = 0; i<=9; i++)
     {printf("& %lf \t",qnorm(v[i]));}
  return 0;
}
```

A.3. Distribución *ji-cuadrada*

A.3.1. Función: double pchisq(double x,int k);

1. Para $k = 1$ y $x > 0$, como $\Gamma(1/2) = \sqrt{\pi}$, y cambiando de variable $z = u/2$,

$$P[X_1 > x] = \frac{1}{\sqrt{\pi}} \int_{x/2}^{\infty} \frac{e^{-z}}{\sqrt{z}} dz = \frac{2}{\sqrt{2\pi}} \int_{\sqrt{x}}^{\infty} e^{-u^2/2} du = 2(1 - \Phi(\sqrt{x}))$$
$$\therefore \quad F_1(x) = 2\ \Phi(\sqrt{x}) - 1$$

2. Para $k = 2$ y $x > 0$, como $\Gamma(1) = 1$,

$$P[X_2 > x] = \int_{x/2}^{\infty} e^{-z}\ dz = e^{-(x/2)}$$
$$\therefore \quad F_2(x) = 1 - e^{-(x/2)}$$

Esta función genera los valores de la función de distribución acumulada F_{χ^2} de la variable aleatoria con distribución *ji-cuadrada* con k grados de libertad.
Para $x > 0$, haciendo el siguiente cambio de variable $z = y/2$,

$$1 - F_{\chi^2} = \int_x^{\infty} f_{\chi^2}(y) dy = \frac{2^{-k/2}}{\Gamma(k/2)} \int_x^{\infty} y^{(k/2)-1} e^{-y/2} dy = \frac{1}{\Gamma(k/2)} \int_{x/2}^{\infty} z^{(k/2)-1} e^{-z} dz$$

Para $k \geq 3$ (integrando por partes: $u = z^{(k/2)-1},\ dv = e^{-z}\ dz$) obtenemos la siguiente función recursiva: $I_n = g(n) I_{n-1}$ con $n = k/2$,

$$\frac{1}{\Gamma(k/2)} \int_{x/2}^{\infty} z^{(k/2)-1} e^{-z} dz = \frac{e^{-(x/2)}\ (x/2)^{(k/2)-1}}{\Gamma(k/2)} + \frac{1}{\Gamma[(k/2)-1]} \int_{x/2}^{\infty} z^{(k/2)-2}\ e^{-z}\ dz$$

Con esta expresión recursiva obtenemos a continuación los valores $F_{\chi^2}(x)$

```
double pchisq(double x, int k)  // x >= 0,  FChi(x,k) = F_k(x)=P[X_k<=x] = y.
{ int d;
  double j,In,y,Gama;
   d = k-(k/2)*2;
        if(d == 0)
           {j = 1.0; In = 1.0;  Gamas = 1.0; }// k es par
        else
           {j = 0.5; In = 2*(1-pnorm(sqrt(x)))/exp(-x/2.0);  Gamas = sqrt(1.0/3.141593); }
   for(double n = j+0.0; n<=(k/2.0)-1.0; n = n+1.0)
            { Gama = n*Gama;
                    +((((n-1)*In)/n)) ;
             Gamas = (n*Gamas)/(n-1);};
                In = (pow((x/2.0),n)/Gama)+In};
                      };
             y=1-(exp(-x/2.0)*In);
   return y;
 }
```

A.3.2. Función: double qchisq(double y,int k)

El siguiente programa genera los valores de la función $F_{\chi^2}^{-1}$ (Tabla 3), correspondiente a la distribución *ji-cuadrada* con k grados de libertad.

```
double qchisq(double y, int k)  /* x >= 0,  FChi(x,k) = F_k(x)=P[X_k<=x] = y, si y solo si  x
{ double a,b,e,x;
 if(k >= 3)
    {  a = 0.0;  b = 150.0;
     do{e = (b-a)/2;  x = a+e;
            if(pchisq(x,k) > y)  {b = x;}
            else
               {if(pchisq(x,k) < y)  {a = x;}
        else{e = 0.0000001;};
               }
      }  while(e > 0.0000001);
    return x;
    }
```

A.3.3. Tabla de la distribución *ji-cuadrada*

El siguiente programa genera la Tabla 3.

```
double pnorm(double x);
double pchisq(double x, int k);
double qchisq(double x, int k);
int main( )  /* TABLA 3 */
{ double v[]={.005,.010,.025,.050,.100,.500,.900,.950,.975,.990,.995 };
    for (int k=1;k<=30;k++)
      {for (int i = 0; i<=10; i++)
     {printf("& %lf \t",qchisq(v[i],k));} ;printf("\\\\ \n");};
  return 0;
}
```

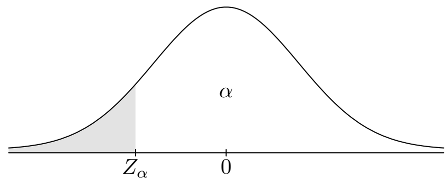


Tabla 1. Valores $1-\alpha$ de la función de distribución acumulada Φ, de la V.A. Z normal estándar.

$$\Phi(z_\alpha) = P[Z \le z_\alpha] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_\alpha} e^{-u^2/2} du = 1 - \alpha$$

z_α	-.00	-.01	-.02	-.03	-.04	-.05	-.06	-.07	-.08	-.09
-3.9	0.000048	0.000046	0.000044	0.000042	0.000041	0.000039	0.000037	0.000036	0.000034	0.000033
-3.8	0.000072	0.000069	0.000067	0.000064	0.000062	0.000059	0.000057	0.000054	0.000052	0.000050
-3.7	0.000108	0.000104	0.000100	0.000096	0.000092	0.000088	0.000085	0.000082	0.000078	0.000075
-3.6	0.000159	0.000153	0.000147	0.000142	0.000136	0.000131	0.000126	0.000121	0.000117	0.000112
-3.5	0.000233	0.000224	0.000216	0.000208	0.000200	0.000193	0.000185	0.000178	0.000172	0.000165
-3.4	0.000337	0.000325	0.000313	0.000302	0.000291	0.000280	0.000270	0.000260	0.000251	0.000242
-3.3	0.000483	0.000466	0.000450	0.000434	0.000419	0.000404	0.000390	0.000376	0.000362	0.000349
-3.2	0.000687	0.000664	0.000641	0.000619	0.000598	0.000577	0.000557	0.000538	0.000519	0.000501
-3.1	0.000968	0.000935	0.000904	0.000874	0.000845	0.000816	0.000789	0.000762	0.000736	0.000711
-3.0	0.001350	0.001306	0.001264	0.001223	0.001183	0.001144	0.001107	0.001070	0.001035	0.001001
-2.9	0.001866	0.001807	0.001750	0.001695	0.001641	0.001589	0.001538	0.001489	0.001441	0.001395
-2.8	0.002555	0.002477	0.002401	0.002327	0.002256	0.002186	0.002118	0.002052	0.001988	0.001926
-2.7	0.003467	0.003364	0.003264	0.003167	0.003072	0.002980	0.002890	0.002803	0.002718	0.002635
-2.6	0.004661	0.004527	0.004396	0.004269	0.004145	0.004025	0.003907	0.003793	0.003681	0.003573
-2.5	0.006210	0.006037	0.005868	0.005703	0.005543	0.005386	0.005234	0.005085	0.004940	0.004799
-2.4	0.008198	0.007976	0.007760	0.007549	0.007344	0.007143	0.006947	0.006756	0.006569	0.006387
-2.3	0.010724	0.010444	0.010170	0.009903	0.009642	0.009387	0.009137	0.008894	0.008656	0.008424
-2.2	0.013903	0.013553	0.013209	0.012874	0.012545	0.012224	0.011911	0.011604	0.011304	0.011011
-2.1	0.017864	0.017429	0.017003	0.016586	0.016177	0.015778	0.015386	0.015003	0.014629	0.014262
-2.0	0.022750	0.022216	0.021692	0.021178	0.020675	0.020182	0.019699	0.019226	0.018763	0.018309
-1.9	0.028717	0.028067	0.027429	0.026803	0.026190	0.025588	0.024998	0.024419	0.023852	0.023295
-1.8	0.035930	0.035148	0.034380	0.033625	0.032884	0.032157	0.031443	0.030742	0.030054	0.029379
-1.7	0.044565	0.043633	0.042716	0.041815	0.040930	0.040059	0.039204	0.038364	0.037538	0.036727
-1.6	0.054799	0.053699	0.052616	0.051551	0.050503	0.049471	0.048457	0.047460	0.046479	0.045514
-1.5	0.066807	0.065522	0.064255	0.063008	0.061780	0.060571	0.059380	0.058208	0.057053	0.055917
-1.4	0.080757	0.079270	0.077804	0.076359	0.074934	0.073529	0.072145	0.070781	0.069437	0.068112
-1.3	0.096800	0.095098	0.093418	0.091759	0.090123	0.088508	0.086915	0.085343	0.083793	0.082264
-1.2	0.115070	0.113139	0.111232	0.109349	0.107488	0.105650	0.103835	0.102042	0.100273	0.098525
-1.1	0.135666	0.133500	0.131357	0.129238	0.127143	0.125072	0.123024	0.121000	0.119000	0.117023
-1.0	0.158655	0.156248	0.153864	0.151505	0.149170	0.146859	0.144572	0.142310	0.140071	0.137857
-0.9	0.184060	0.181411	0.178786	0.176186	0.173609	0.171056	0.168528	0.166023	0.163543	0.161087
-0.8	0.211855	0.208970	0.206108	0.203269	0.200454	0.197663	0.194895	0.192150	0.189430	0.186733
-0.7	0.241964	0.238852	0.235762	0.232695	0.229650	0.226627	0.223627	0.220650	0.217695	0.214764
-0.6	0.274253	0.270931	0.267629	0.264347	0.261086	0.257846	0.254627	0.251429	0.248252	0.245097
-0.5	0.308538	0.305026	0.301532	0.298056	0.294599	0.291160	0.287740	0.284339	0.280957	0.277595
-0.4	0.344578	0.340903	0.337243	0.333598	0.329969	0.326355	0.322758	0.319178	0.315614	0.312067
-0.3	0.382089	0.378280	0.374484	0.370700	0.366928	0.363169	0.359424	0.355691	0.351973	0.348268
-0.2	0.420740	0.416834	0.412936	0.409046	0.405165	0.401294	0.397432	0.393580	0.389739	0.385908
-0.1	0.460172	0.456205	0.452242	0.448283	0.444330	0.440382	0.436441	0.432505	0.428576	0.424655
-0.0	0.500000	0.496011	0.492022	0.488034	0.484047	0.480061	0.476078	0.472097	0.468119	0.464144

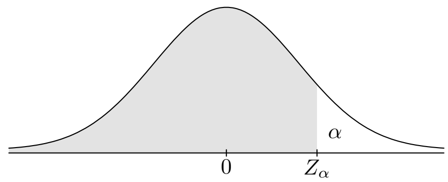


Tabla 1. (Continuación) Valores $1-\alpha$ de la función de distribución acumulada Φ, de la V.A. Z normal estándar.

$$\Phi(z_\alpha) = P[Z \leq z_\alpha] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_\alpha} e^{-u^2/2} du = 1-\alpha$$

z_α	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.500000	0.503989	0.507978	0.511966	0.515953	0.519939	0.523922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555670	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621720	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802337	0.805105	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823814	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.879000	0.881000	0.882977
1.2	0.884930	0.886861	0.888768	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903200	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935745	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959070	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965620	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999534	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999651
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999822	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967

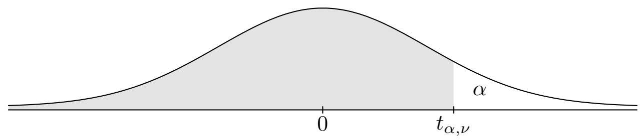


Tabla 2. Puntos porcentuales $t_{\alpha,\nu}$ de la distribución t.

$$F_{\nu}(t_{\alpha,\nu}) = P[T_\nu \leq t_{\alpha,\nu}] = \frac{\Gamma[(k+1)/2]}{\sqrt{k\pi}\ \Gamma(k/2)} \int_{-\infty}^{t_{\alpha,\nu}} \left(1+\frac{u^2}{k}\right)^{-(k+1)/2} du = 1-\alpha$$

$$F_{\nu}^{-1}(1-\alpha) = t_{\alpha,\nu}$$

$1-\alpha$	.70	.75	.90	.95	.975	.99	.995	.9975	.999	.9995
ν \ α	.30	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	0.726543	1.000000	3.077685	6.313758	12.70623	31.82069	63.65744	127.3241	318.3263	636.6894
2	0.617213	0.816497	1.885618	2.919986	4.302653	6.964557	9.924843	14.08905	22.32712	31.59905
3	0.584383	0.764894	1.637735	2.353373	3.182440	4.540720	5.840940	7.453375	10.21472	12.92445
4	0.568647	0.740690	1.533213	2.131853	2.776442	3.746939	4.604101	5.597563	7.173185	8.610296
5	0.559435	0.726690	1.475878	2.015047	2.570581	3.364935	4.032145	4.773359	5.893507	6.868982
6	0.553389	0.717554	1.439753	1.943178	2.446909	3.142672	3.707438	4.316835	5.207624	5.958815
7	0.549116	0.711145	1.414919	1.894579	2.364626	2.997961	3.499498	4.029360	4.785337	5.407991
8	0.545931	0.706396	1.396818	1.859541	2.306013	2.896452	3.355379	3.832521	4.500799	5.041304
9	0.543489	0.702715	1.383028	1.833105	2.262163	2.821436	3.249846	3.689680	4.296846	4.780989
10	0.541525	0.699816	1.372175	1.812468	2.228136	2.763777	3.169279	3.581400	4.143705	4.586897
11	0.539942	0.697451	1.363440	1.795893	2.200994	2.718077	3.105822	3.496618	4.024725	4.437037
12	0.538626	0.695486	1.356211	1.782293	2.178812	2.680998	3.054533	3.428450	3.929625	4.317789
13	0.537500	0.693827	1.350164	1.770926	2.160368	2.650309	3.012285	3.372488	3.852015	4.220896
14	0.536547	0.692415	1.345034	1.761312	2.144785	2.624502	2.976847	3.325701	3.787394	4.140463
15	0.535727	0.691195	1.340609	1.753054	2.131453	2.602491	2.946711	3.286047	3.732862	4.072828
16	0.535002	0.690126	1.336756	1.745882	2.119913	2.583494	2.920790	3.252001	3.686152	4.014997
17	0.534372	0.689192	1.333380	1.739607	2.109823	2.566938	2.898245	3.222456	3.645792	3.965178
18	0.533819	0.688372	1.330385	1.734056	2.100916	2.552385	2.878447	3.196573	3.610487	3.921652
19	0.533323	0.687628	1.327734	1.729136	2.093019	2.539492	2.860937	3.173742	3.579416	3.883448
20	0.532866	0.686960	1.325350	1.724710	2.085962	2.527971	2.845335	3.153410	3.551817	3.849516
21	0.532446	0.686350	1.323195	1.720743	2.079611	2.517653	2.831373	3.135214	3.527174	3.819323
22	0.532084	0.685797	1.321230	1.717138	2.073870	2.508326	2.818747	3.118830	3.504992	3.792124
23	0.531740	0.685301	1.319456	1.713877	2.068663	2.499876	2.807341	3.104010	3.484983	3.767672
24	0.531435	0.684843	1.317835	1.710882	2.063894	2.492151	2.796946	3.090506	3.466768	3.745394
25	0.531149	0.684423	1.316347	1.708136	2.059546	2.485113	2.787447	3.078203	3.450212	3.725195
26	0.530901	0.684042	1.314974	1.705618	2.055521	2.478628	2.778711	3.066912	3.434992	3.706617
27	0.530653	0.683680	1.313696	1.703291	2.051840	2.472658	2.770681	3.056536	3.421049	3.689642
28	0.530424	0.683355	1.312532	1.701136	2.048407	2.467146	2.763262	3.046923	3.408155	3.673906
29	0.530214	0.683050	1.311426	1.699133	2.045221	2.462015	2.756395	3.038054	3.396254	3.659449
30	0.530024	0.682764	1.310415	1.697264	2.042265	2.457266	2.749987	3.029795	3.385191	3.645964
∞	0.524401	0.674490	1.281552	1.644854	1.959964	2.326348	2.575829	2.807034	3.090232	3.290527

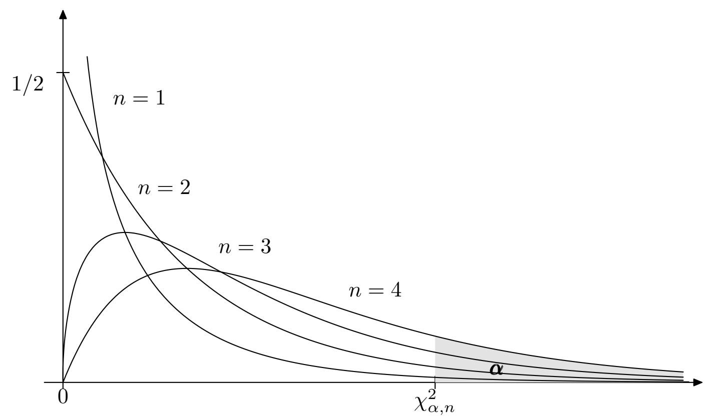


Tabla 3. Puntos porcentuales $\chi^2_{\alpha,n}$ de la distribución $Ji - cuadrada$.

$$F_n(\chi^2_{\alpha,n}) = P[X_n \leq \chi^2_{\alpha,n}] = \int_{-\infty}^{\chi^2_{\alpha,n}} f_n(u)\ du = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^{\chi^2_{\alpha,n}} u^{(n/2)-1} e^{-u/2} du = 1 - \alpha$$

n \ α	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
1	$4(10^{-5})$	$2(10^{-4})$	10^{-3}	$4(10^{-3})$	0.016	0.455	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	1.386	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	2.366	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	3.357	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	4.351	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	5.348	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	6.346	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	7.344	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	8.343	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	9.342	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	10.341	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	11.340	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	12.340	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	13.339	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	14.339	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	15.338	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	16.338	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	17.338	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	18.338	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	19.337	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	20.337	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	21.337	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	22.337	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	23.337	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	24.337	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	25.336	35.563	38.885	41.923	45.642	48.290
27	11.808	12.878	14.573	16.151	18.114	26.336	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	27.336	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	28.336	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	29.336	40.256	43.773	46.979	50.892	53.672

Bibliografía

[1] MILLER Charles D. HEEREN Vern E. HORNSBY E. JOHN *Matemática: Razonamiento y Aplicaciones*, Addison Wesley Longman , México 1999.

[2] ROSS Sheldom M. *A First Course in Probability*, Prentice Hall , New Jersey 1998.

[3] MEYER Paul L. *PROBABILIDAD Y APLICACIONES ESTADÍSTICAS*, Addison Wesley Longman , México 1998.

[4] WACKERLY Dennis D. MENDENHALL Willian SCHEAFFER Richard *ESTDÍSTICA MATEMÁTICA CON APLICACIONES*, THOMSON , México 2002.

[5] HINES William W. MONTGOMERY Douglas C. *PROBABILIDAD Y ESTADÍSTICA PARA INGENIERÍA Y ADMINISTRACIÓN*, CECSA , México 1996.

[6] CHUNG Kai Lai *A COURSE IN PROBABILITY THEORY*, Harcourt, Brace & World, Inc. , USA 1968.

[7] TUCKER Howard G. *A GRADUATE COURSE IN PROBABILITY*, Academic Press, New York San Francisco London 1967.

Estadística y probabilidad
se terminó de imprimir en Marzo de 2025,
en los talleres de la Universidad Autónoma de la
Ciudad de México, San Lorenzo, 290, col. Del Valle,
Alcaldía Benito Juárez, c.p. 03100, Ciudad de México.
El tiraje fue de 500 ejemplares.
Cuidado de la edición: Ángeles Godínez Guevara
Diseño editorial: Sergio Cortés Becerril

Manuel Jesús Tec Canché. Realizó estudios de Licenciatura en la Escuela Superior de Física y Matemáticas (ESFM) del Instituto Politécnico Nacional (IPN); y su Maestría en Ciencias en el Departamento de Matemáticas del Centro de Investigación y de Estudios Avanzados (CINVESTAV) del Instituto Politécnico Nacional (IPN). Como docente ha trabajado en la Universidad Anahuac México Campus Norte; en la Universidad Autónoma Metropolitana (UAM) Unidad Iztapalapa; y actualmente es Profesor Investigador de Tiempo Completo de la Universidad Autónoma de la Ciudad de México (UACM) Plantel San Lorenzo Tezonco.

Estadística y Probabilidad es el resultado de varios años de enseñanza en Universidades; tiene como objetivo apoyar a los estudiantes de Ingeniería y de Modelación matemática de la Universidad Autónoma de la Ciudad de México en su curso de Estadística y probabilidad donde se profundiza en la parte de la Probabilidad ya que es el lenguaje de la Estadística; y donde se utiliza como herramienta el software estadístico libre R.

UACM
Universidad Autónoma
de la Ciudad de México
NADA HUMANO ME ES AJENO

Biblioteca del Estudiante

9 786072 615489