

# STATISTIC STATISTIC

RICARDO



Departamento de matemática y física, FIMGC-USNCH  
*E-mail:* ricardomallqui6@gmail.com  
*URL:* [www.fractales.com](http://www.fractales.com)

Estadística en artistas  
Ricardo Michel Mallqui Baños

Un libro de estadística, basado en código Asymptote y LaTeX.

Bibliografía.

Incluye Índice.

1. Estadística descriptiva 2. Estadística Inferencial



Todos los derechos reservados. Ninguna parte de este libro puede ser reproducido en cualquier forma, o por cualquier medio, sin permiso por escrito del editor.

Departamento de matemática y física, FIMGC-USNCH

*E-mail:* ricardomallqui6@gmail.com

**URL:** [www.fractales.com](http://www.fractales.com)

# Índice general

# Lista de figuras

# Lista de tablas

# Presentación

Este libro se elaboró

Double underlined text Double underlined text

*Slanted underlined* SMALL CAPS UNDERLINED

Parte I

**Estadística descriptiva**

# Capítulo 1

## Preliminares

### 1.1. Conceptos básicos

#### 1.1.1. Parámetros

**Definición 1.1 (Poblacion)** *A statistical population is a data set (usually large, sometimes conceptual) that is our target of interest.*

**Definición 1.2 (Poblacion)** *A statistical population is a data set (usually large, sometimes conceptual) that is our target of interest.*

**Definición 1.3 (Poblacion)** *A statistical population is a data set (usually large, sometimes conceptual) that is our target of interest.*

**Definición 1.4 (Poblacion)** *A statistical population is a data set (usually large, sometimes conceptual) that is our target of interest.*

**Definición 1.5 (Poblacion)** *A statistical population is a data set (usually large, sometimes conceptual) that is our target of interest.*

**Teorema 1.6** *wwwwwwwwwwwwwwwwww*

#### 1.1.2. Variables estadísticas

**Definición 1.7 (Variables cuantitativas)** *Una variable estadística es una característica que puede fluctuar y cuya variación es susceptible de adoptar diferentes valores, los cuales pueden medirse u observarse. Las variables adquieren valor cuando se relacionan con otras variables, es decir, si forman parte de una hipótesis o de una teoría. Existen dos clases de variables: Cualitativas y cuantitativas.*

1. *Cualitativas. Son aquellas variables que están propensos a ser nominadas textualmente.*
  - a) *Nominales. Son características que simplemente nominan y están propensos a ser jerarquizados u ordenados tales como: El estado civil (soltero, casado, divorciado, viudo), Religión (católica, evangélico, judío, etc).*



- b) *Ordinales. Son características que que si están propensos a ser jerarquizados tales como: Nivel de instrucción (inicial, primaria, secundaria, superior).*
- 2. *Cuantitativas. Son aquellas variables que están propensos a ser medidas mediante números ya sean números enteros o reales.*
  - a) *Discretas. Aquellas que solo son medidos mediante números enteros por ejemplo: Número de hijos y número de habitaciones.*
  - b) *Continuas. Aquellas que solo son medidos mediante números reales es decir este incluye a los números racionales e irracionales. Estatura, volumen, peso.*

## 1.2. Organización de datos en tablas de frecuencias

### 1.2.1. Distribución de frecuencias

El uso de tablas de distribución de frecuencias y gráficas como un medio para presentar la información de un conjunto de datos de forma resumida. En grados anteriores ya se ha trabajado con gráficas para variables cuantitativas discretas, por lo que esta será la primera vez que el estudiante trabajará con gráficas que son adecuadas para presentar información de variables cuantitativas continuas.

**Definición 1.8** *La tabulación es un proceso en el cual los datos son ordenados en grupos llamados clases para un análisis más eficaz de estos, los datos podrían estar clasificados mediante una variable cualitativa o cuantitativa en el caso de las variables cualitativas  $Y_i$ , se considera la siguiente Tabla 1.1*

En el caso de variables cuantitativas además si los datos son muy variados, que para se clasificados adecuadamente, necesitan generarse particiones de longitudes semejantes entonces se utiliza el siguiente proceso; el número de las particiones  $r$  se consideran de acuerdo a tres criterios.

1. Criterio del investigador  $r$  no puede ser más de 20 ni menos de 5
2.  $r = \sqrt{n}$  donde  $n$  es el número de datos
3. La regla de Starges que consiste en considerar la fórmula  $r = 3,322 \cdot \log_{10} n$   
Una vez establecido el número de particiones se procede a generar los límites laterales de cada una de las particiones, sea  $L$  la longitud de todo el conjunto es decir  $L = x_{\max} - x_{\min}$  entonces la longitud de las particiones o amplitud intervállica se obtiene con  $l = \frac{L}{r}$

Cuadro 1.1: Combinaciones de los tres segmentos de la seccion aurea.

$Y_i$	$f_i$	$F_i$	$F_i^*$	$h_i$	$H_i$	$H_i$	$h_i \%$	$H_i \%$	$H_i * \%$
2	1	1.00	20.00	0.05	0.05	0.0025	0.25	0.25	0.2500
3	2	3.00	19.00	0.10	0.15	0.0050	0.50	0.50	0.7500
4	5	8.00	17.00	0.25	0.40	0.0125	1.25	1.25	2.0000
5	7	15.00	12.00	0.35	0.75	0.0175	1.75	1.75	3.7500
6	4	19.00	5.00	0.20	0.95	0.0100	1.00	1.00	4.7500
7	1	20.00	1.00	0.05	1.00	0.0025	0.25	0.25	5.0000
$\Sigma$	20						5		

Tenga en cuenta que  $n$  es el número de datos, es decir

$$n = f_1 + f_2 + \dots + f_r = \sum_{i=1}^r f_i$$

donde  $f_i$  es número de datos en la partición  $X_i$ , una de las  $r$  particiones del conjunto total de datos.

1. Las frecuencias absolutas  $f_i$  indican el número de datos con la característica  $X_i$ .
2. Las frecuencias absolutas acumuladas menor que  $F_i$  se obtienen mediante la fórmula

$$F_m = f_1 + f_2 + \dots + f_m = \sum_{i=1}^m f_i.$$

3. Las frecuencias absolutas acumuladas mayor que  $F_i^*$  obedecen a la fórmula

$$\begin{aligned} F_m^* &= f_m + f_{m+1} + \dots + f_r \\ &= \sum_{i=m}^r f_i = n - \sum_{i=1}^{m-1} f_i = n - (f_1 + f_2 + \dots + f_{m-1}) \end{aligned}$$

4. Las frecuencias absolutas relativas se obtienen mediante la fórmula

$$h_m = \frac{f_m}{n}$$

5. Las frecuencias absolutas relativas menor que obedecen a la fórmula

$$H_m = \frac{f_m}{n}$$

6. Las frecuencias absolutas relativas mayor que

$$H_m^* = \frac{F_m}{n}$$

7. Las frecuencias absolutas relativas porcentuales  $h_i \% = 100 \cdot h_i$

8. Las frecuencias absolutas relativas menor que porcentuales  $H_i \% = 100 \cdot H_i$
9. Las frecuencias absolutas relativas mayor que porcentuales  $H_i^* \% = 100 \cdot H_i^*$
10.  $Y_i$  marca de clase o punto medio de la clase  $i$

### 1.2.2. Distribución de frecuencias continuas

## 1.3. Gráficos estadísticos

### 1.3.1. Histograma de frecuencias

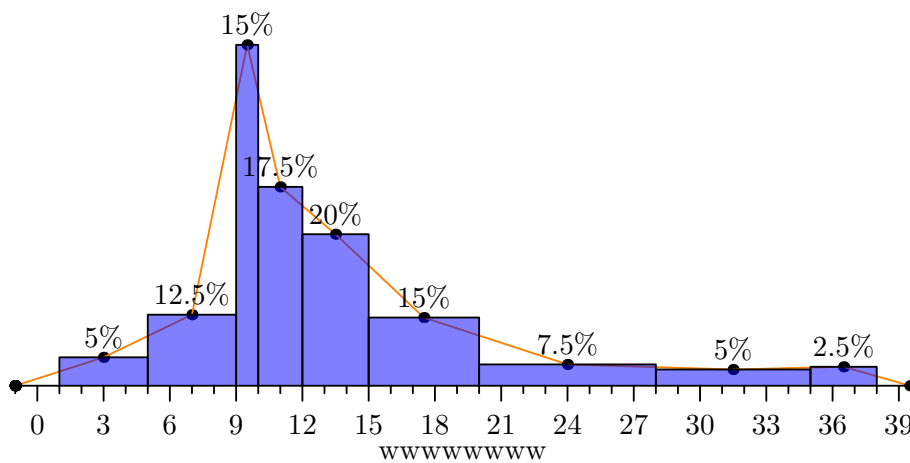


Figura 1.1: Histograma de frecuencias

### 1.3.2. Gráficos circulares

Si se tiene  $n$  clases entonces el sector circular tendra una radio de  $360 \cdot h_i$

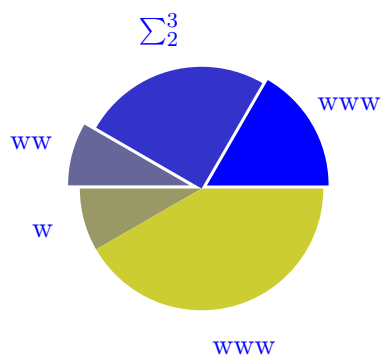


Figura 1.2: Circular

### 1.3.3. Diagrama de barras

Este tipo de diagramas representan aquellos datos cuantitativos, clasificados en clases

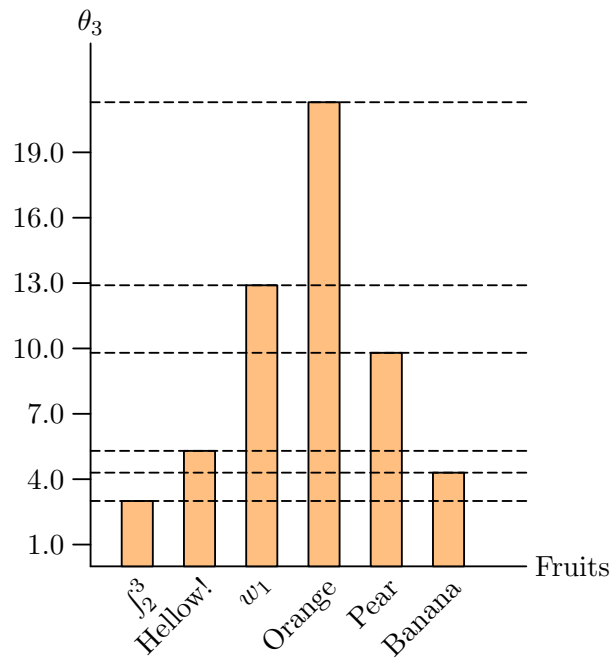


Figura 1.3: Diagrama de barras

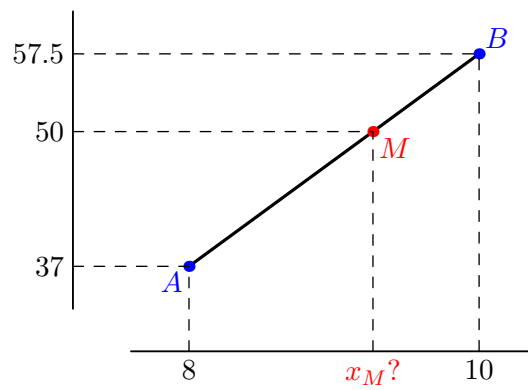


Figura 1.4: Meadiana

## Capítulo 2

# Medidas estadísticas de variables cuantitativas

### 2.1. Medidas de tendencia central

Son aquellas medidas que buscan un dato representativo central de un conjunto de datos tales como la media, la moda y la mediana.

**Definición 2.1 (Datos agrupados y no agrupados)** *La principal diferencia entre los datos agrupados y los no agrupados es que los agrupados están clasificados según un criterio y los no agrupados se encuentran en el mismo formato que cuando se*

**Comentario 2.1 ()**

.

Cuadro 2.1: Combinaciones de los tres segmentos de la sección anterior.

Clase	$Y_i$	$f_i$	$F_i$	...	$H_i^* \%$
$[y_1, y_2)$	$y_1$	$f_1$	...	...	$H_1^* \%$
$[y_2, y_3)$	$y_2$	$f_2$	...	...	$H_1^* \%$
$[y_3, y_4)$	$y_3$	$f_3$	...	...	$H_1^* \%$
$\vdots$	$\vdots$	$\vdots$	...	...	$\vdots$
$[y_{r-1}, y_r]$	$y_r$	$f_r$	...	...	$H_1^* \%$

#### 2.1.1. Media o promedio

1. Datos no agrupados.  $\bar{x} = \frac{\sum_1^n x_i}{n}$  donde  $n$  es el número de datos.
2. Datos agrupados.  $\bar{x} = \frac{\sum_1^m f_i y_i}{n}$  donde  $m < n$ ,  $m$  número de clases y  $n$  es el número de datos.

**2.1.2. Mediana**

**2.1.3. Moda**

**2.2. Medidas de dispersión**

**2.3. Medidas de variación**

## Capítulo 3

# Variables estadísticas bidimensionales

When we put (vertically) large expressions inside of parentheses (or brackets, or curly braces, etc.), the parentheses don't resize to fit the expression and instead remain relatively small. For instance,

$$f(x) = \pi\left(\frac{\sqrt{x}}{x-1}\right)$$

comes out as

When we put (vertically) large expressions inside of parentheses (or brackets, or curly braces, etc.), the parentheses don't resize to fit the expression and instead remain relatively small. For instance,

$$f(x) = \pi\left(\frac{\sqrt{x}}{x-1}\right)$$

comes out as

When we put (vertically) large expressions inside of parentheses (or brackets, or curly braces, etc.), the parentheses don't resize to fit the expression and instead remain relatively small. For instance,

$$f(x) = \pi\left(\frac{\sqrt{x}}{x-1}\right)$$

comes out as

Cuadro 3.1: Combinaciones de los tres segmentos de la seccion aurea.

		$Y_i$	$f_i$	$F_i$	$F_i^*$	$h_i$	$H_i$	$H_i$	$h_i$ %	$H_i$ %	$H_i^*$ %
2	1	1.0000	20.0000	0.0500	0.0500	0.0025	0.2500	0.2500	0.2500		
3	2	3.0000	19.0000	0.1000	0.1500	0.0050	0.5000	0.5000	0.7500		
4	5	8.0000	17.0000	0.2500	0.4000	0.0125	1.2500	1.2500	2.0000		
Continúa en la proxima página											

Cuadro 3.1 – continua de la página anterior

		$Y_i$	$f_i$	$F_i$	$F_i^*$	$h_i$	$H_i$	$H_i$	$h_i$ %	$H_i$ %	$H_i *$ %	
5	7	15.0000	12.0000	0.3500	0.7500	0.0175	1.7500	1.7500	3.7500			
6	4	19.0000	5.0000	0.2000	0.9500	0.0100	1.0000	1.0000	4.7500			
7	1	20.0000	1.0000	0.0500	1.0000	0.0025	0.2500	0.2500	5.0000			
20											5	

Cuadro 3.2: www.

Variables	$y_1$	$y_2$	$\dots$	$y_j$	$\dots$	$y_n$	Total
$x_1$	$f_{11}$	$f_{12}$	$\dots$	$f_{1j}$	$\dots$	$f_{1e}$	$f_1$
$x_2$	$f_{21}$	$f_{22}$	$\dots$	$f_{2j}$	$\dots$	$f_{2e}$	$f_2$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$x_i$	$f_{i1}$	$f_{i2}$	$\dots$	$f_{ij}$	$\dots$	$f_{ie}$	$f_j$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$x_k$	$f_{k1}$	$f_{k2}$	$\dots$	$f_{kj}$	$\dots$	$f_{ke}$	$f_k$
Total	$f_1$	$f_2$	$\dots$	$f_j$	$\dots$	$f_n$	$n = \sum_1^n \sum_1^n f_{ij}$



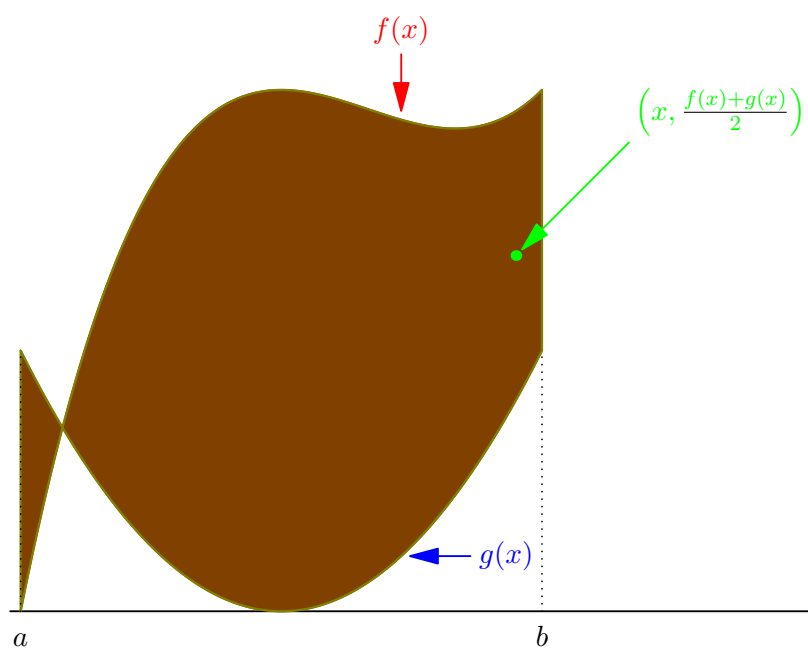


Figura 3.1:  $f(x)$   $\left(x, \frac{f(x)+g(x)}{2}\right)$

**Parte II**

**Cálculo de probabilidades**

## Capítulo 4

## Variables aleatorias

**Definición 4.1 (Experimento aleatorio)** *El porceso en el cual .....  
Se donata con  $\epsilon$*

**Definición 4.2 (Espacio muestral)** *El porceso en el cual ..... Se donda con  $\epsilon$*

**Definición 4.3 (Suceso)** *El porceso en el cual ..... Se donata con  $\epsilon$*

**Definición 4.4 (Evento)** *El porceso en el cual ..... Se donata con  $\epsilon$*

Definición 4.5 (title) ~~~~~

**Comentario 4.1 (title)** *xxxxxxxxxxxxxxxxxxxx*

**Teorema 4.6 (title)** wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww

**Proposición 4.6.1 (title)** *wwwwwwwwwwwwww*

## Capítulo 5

WWW

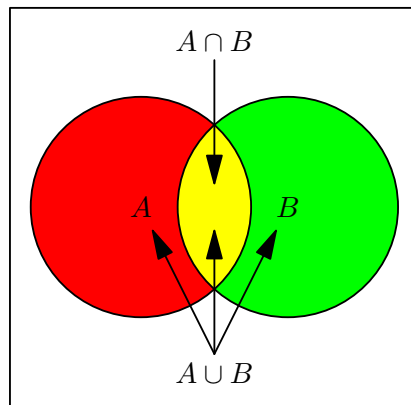


Figura 5.1:

## Capítulo 6

# Variables aleatorias discretas

## Capítulo 7

# Variables aleatorias continuas

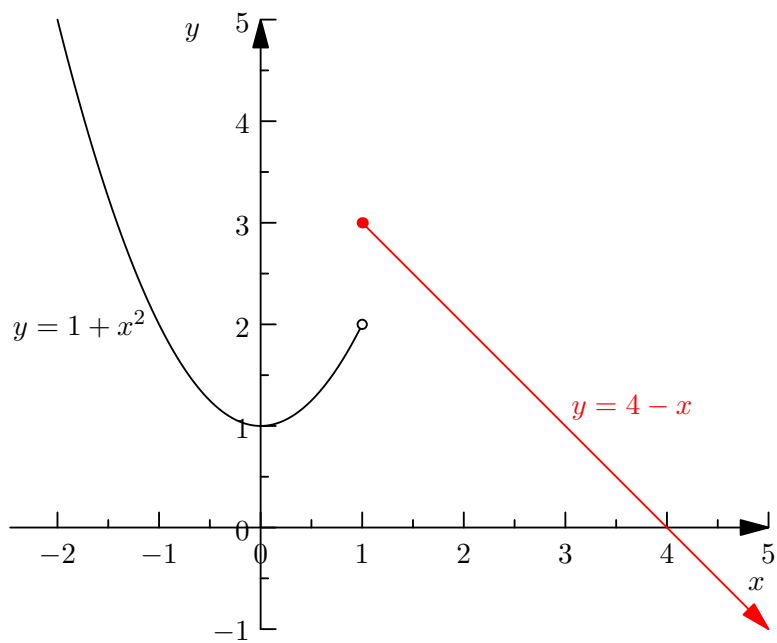


Figura 7.1: www2

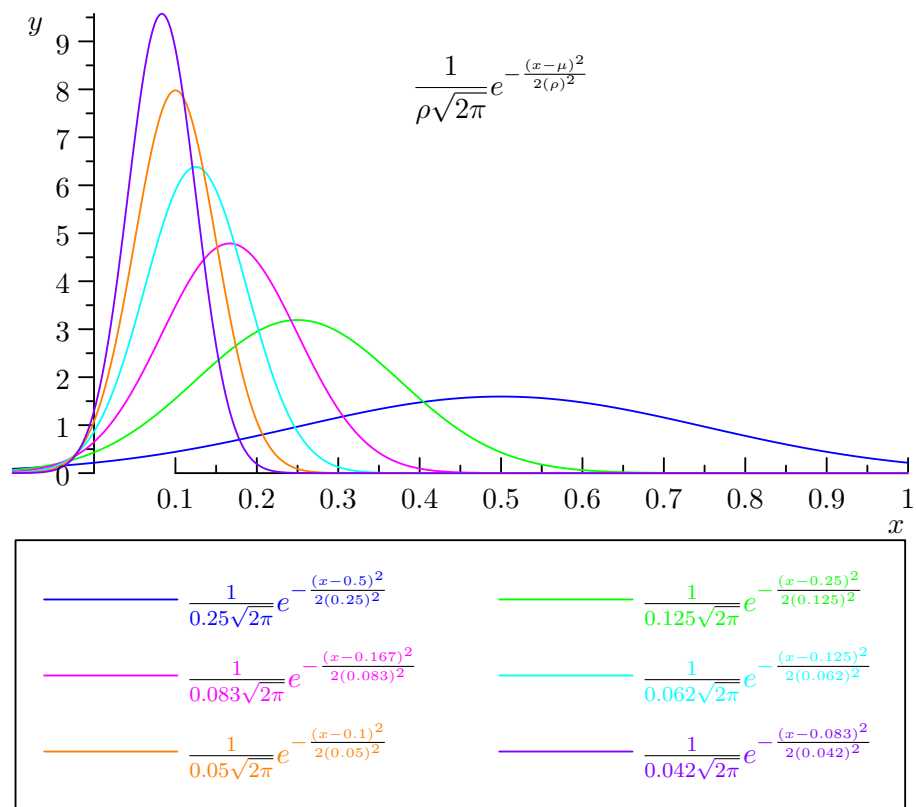


Figura 7.2: wwwwww

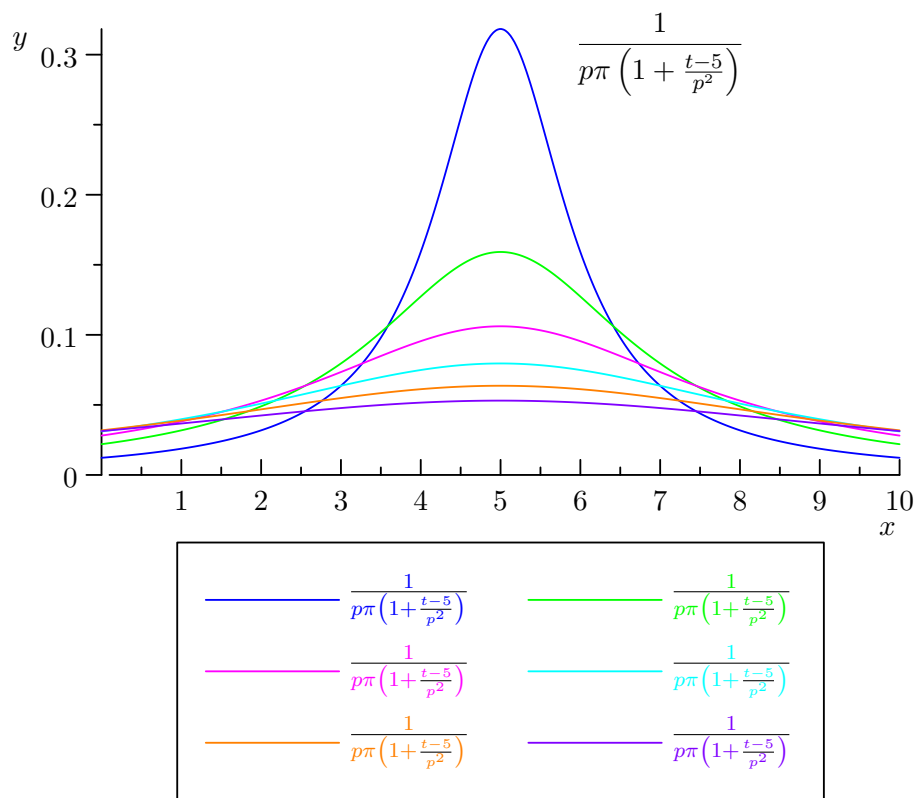


Figura 7.3: wwwwww



**Parte III**

**Inferencial estadística**

La inferencia estadística es primordialmente de naturaleza inductiva y llega a generalizar respecto de las características de una población valiéndose de observaciones empíricas de la muestra.

Al utilizar estadísticas muestrales para estudiar un parámetro de la población es muy normal que ambos sean diferentes y la igualdad entre ambos sea mera coincidencia. La diferencia entre la estadística muestral y el correspondiente parámetro de la población se suele llamar error de estimación. Solo conoceríamos dicho error si se conociera el parámetro poblacional que por lo general se desconoce. La única forma de tener alguna certeza al respecto es hacer todas las observaciones posibles del total de la población; en la mayoría de las aplicaciones prácticas es imposible o impracticable.

## Capítulo 8

# Distribuciones muestrales

La inferencia estadística es primordialmente de naturaleza inductiva y llega a generalizar respecto de las características de una población valiéndose de observaciones empíricas de la muestra.

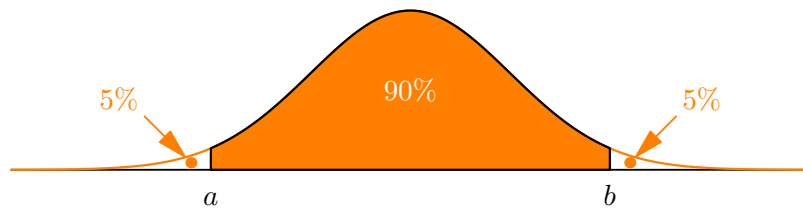


Figura 8.1: Normal

(?, ?). “Geometry and the Imagination” this environment shares the counter of the previously defined thm environment. (?, ?) (?, ?)

## Apéndice A

# Sistemas de coordenadas