

A photograph of two lion cubs in a savanna setting. One cub is on the left, facing right, and the other is on the right, facing left. They are both standing on their hind legs and reaching out with their front paws towards each other, as if playing. The background is a blurred green field with some tall grass. The text "AST0212 – 2016-1" is overlaid in the center of the image.

AST0212 – 2016-1

Introducción al análisis de datos

Instituto de Astrofísica

Facultad de Física

Pontificia Universidad Católica de Chile

Equipo docente:

Profesor: Alejandro Clocchiatti

Ayudantes:

Francisco Aros (TM6)

Nicolás Castro (TL4)

TM6: Tutoría del martes en módulo 6

TL4: Tutoría del lunes en módulo 4

Nuestro Semestre 2016-1

AST0212

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
6 Mar 2016 Semana 1	7	8	9	10	11 C1 ✓	12
13 Semana 2	14 TL1	15 TM1	16	17	18 C2	19
20 Semana 3	21 TL2	22 TM2	23	24	25 Feriado	26
27 Semana 4	28 TL3	29 TM3	30	31	1 Apr C3	2
3 Semana 5	4 TL4	5 TM4	6	7	8 C4	9
10 Semana 6	11 TL5	12 TM5	13	14	15 C5	16
17 Semana 7	18 TL6	19 TM6	20	21	22 C6	23
24 Semana 8	25 TL7	26 TM7	27	28	29 C7	30
1 May Semana 9	2 TL8	3 TM8	4	5	6 C8	7
8 Semana 10	9 TL9	10 TM9	11	12	13 C9	14
15 Semana 11	16 TL10	17 TM10	18	19	20 C10	21
22 Semana 12	23 TL11	24 TM11	25	26	27 C11	28
29 Semana 13	30 TL12	31 TM12	1 Jun	2	3 Feriado	4
5 Semana 14	6 TL13	7 TM13	8	9	10 C12	11
12 Semana 15	13 TL14	14 TM14	15	16	17 C13	18
19	20	21	22	23	24	25
		28			1 Jul	2
		5			8 Notas	9

← Control 1

Tutorías día lunes
Módulo 4:
Nicolás Castro

Tutorías día martes
Módulo 6:
Francisco Aros

Clase previa (Clase 1):

1. Revisión de temas de Clase 0
 1. Organización de reuniones semanales
 2. Evaluación del trabajo de los alumnos (¿?)
2. Datos para Tarea 1
 1. ¿Status de toma de datos?
 2. “Fake” data y ejemplo de uso de herramientas Linux
3. Vueltas de tuerca sobre la Tarea 1
 1. ¿Cuál es la mejor balanza?

Esto nos quedó pendiente para la Clase 2 (hoy).

Esta clase (Clase 2):

1. Temas pendientes de la Clase 1

1. Datos para Tarea 1

1. ¿Status de toma de datos?
2. “Fake” data y ejemplo de uso de herramientas Linux

¿Datos listos?

¿Tablas listas?

Sistemas Linux: Seguir practicando. ¡Es el futuro!

2. Vueltas de tuerca sobre la Tarea 1

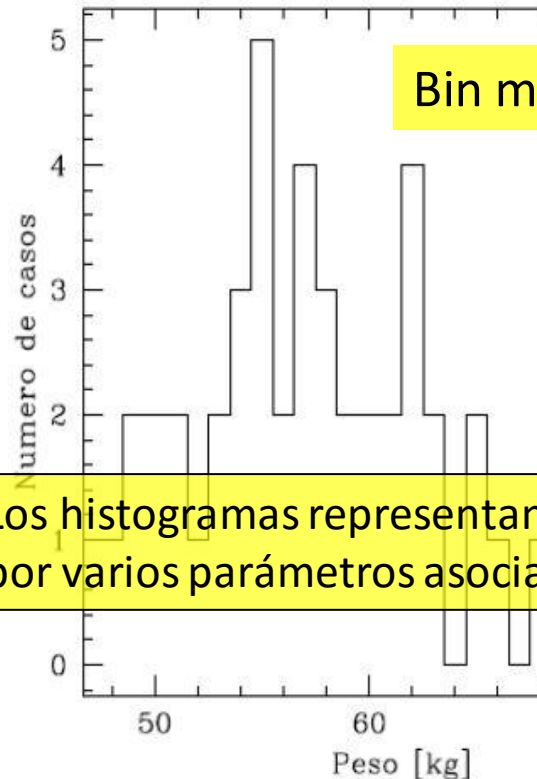
1. ¿Cómo visualizar fácilmente cientos de datos?
2. ¿Cuál es la mejor balanza?

Temas del día: Histogramas, errores aleatorios y sistemáticos.

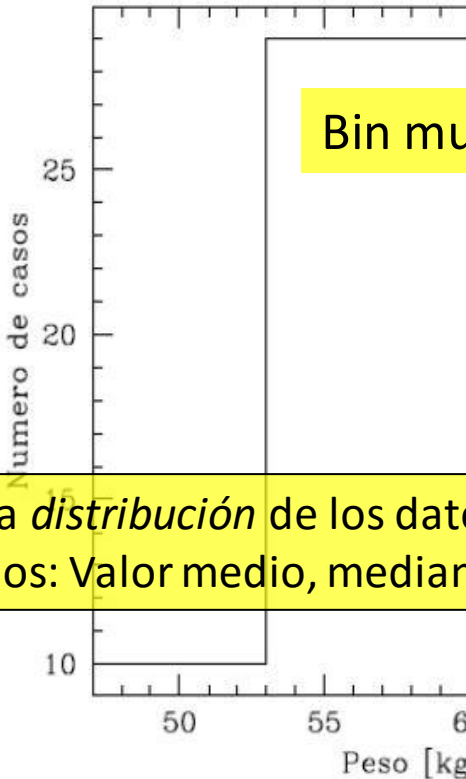
Histogramas

El objetivo de los histogramas es proporcionar una visión rápida y compacta de una gran cantidad de datos directamente comparables y ver como se organizan de acuerdo a su valor. Hay algo de arte en esto de construir un histograma:

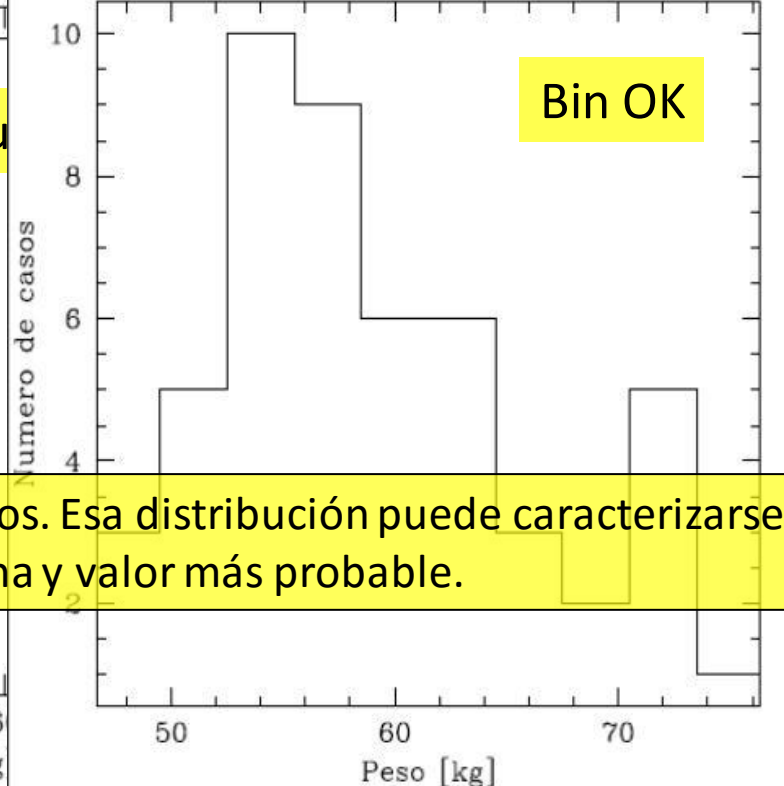
cristobal-moya.dat ; fem ; $N_T = 50$



cristobal-moya.dat ; fem ; N_T

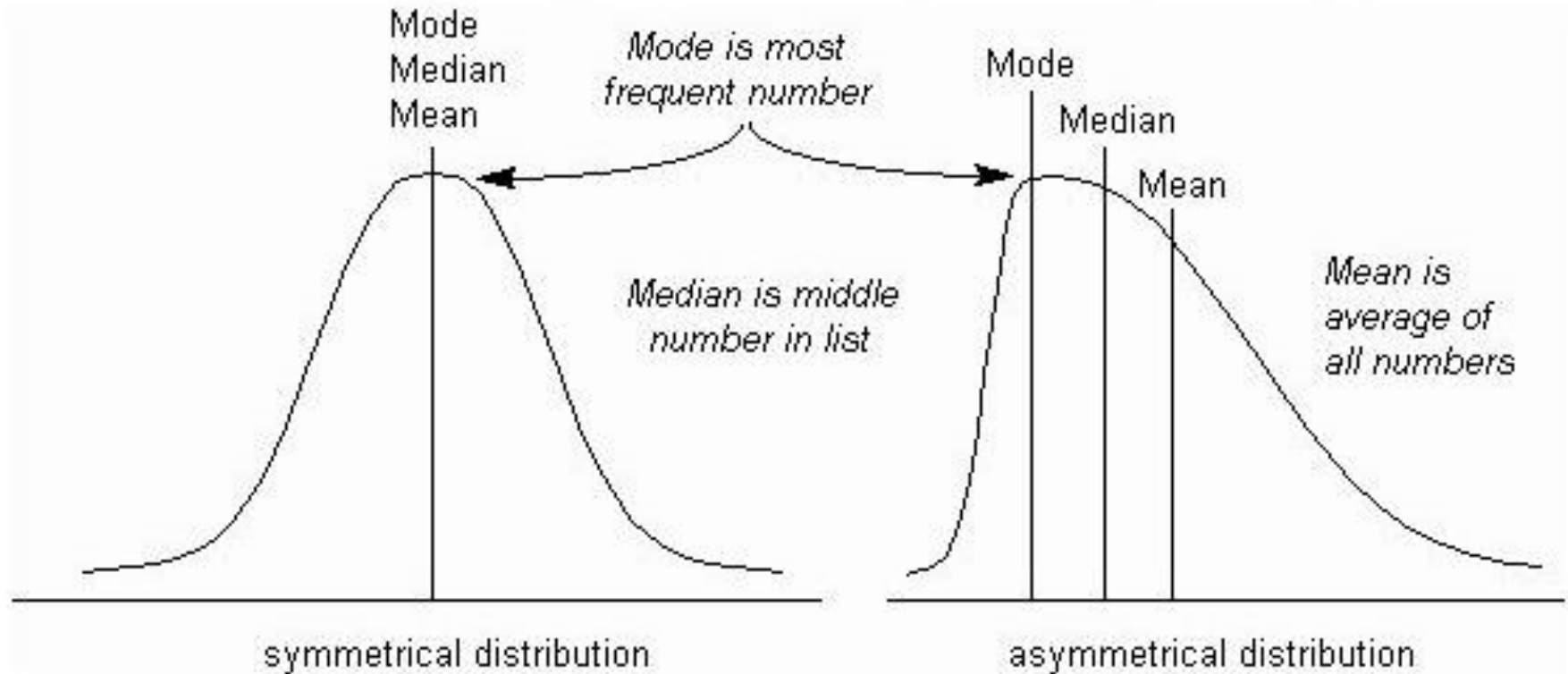


cristobal-moya.dat ; fem ; $N_T = 50$; Bin = 3 kg



Los histogramas representan la *distribución* de los datos. Esa distribución puede caracterizarse por varios parámetros asociados: Valor medio, mediana y valor más probable.

Histogramas: Media, mediana y moda

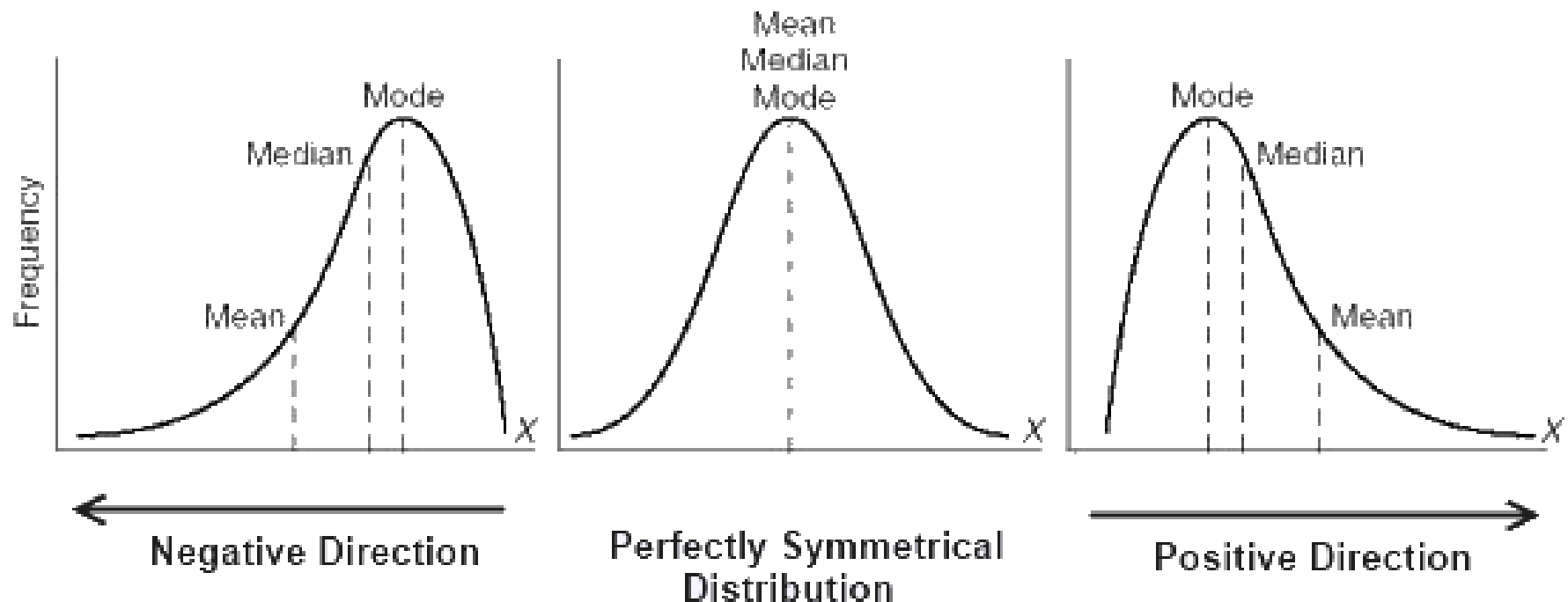


Histogramas: Media, mediana y moda

(a) Negatively skewed

(b) Normal (no skew)

(c) Positively skewed



Las distribuciones, y sus histogramas, pueden ser simétricas o estar sesgados hacia un lado u otro. Ésto se define arbitrariamente como sesgo positivo (a la derecha) o negativo (a la izquierda).

Histogramas

Calcular la mediana y la moda requiere contar valores.

Calcular la media requiere aplicar fórmulas:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

“ i ” recorre todos los datos, del primero al último (1 a N).

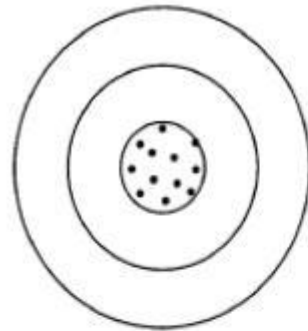
$$\overline{x_g} = \frac{1}{N} \sum_{j=1}^M n_j \bar{x}_j$$

“ j ” recorre el número de bins, del primero al último (1 a M). \bar{x}_j es el valor medio del j -ésimo bin.

Pregunta para pensar en casa: ¿Son consistentes estas definiciones?
¿Dan el mismo valor medio?

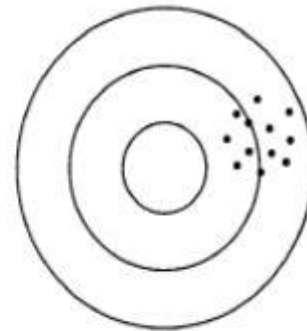
¡Fin de la clase!
Pasamos a toma de datos para
finalizar la Tarea 1

Errores aleatorios y sistemáticos



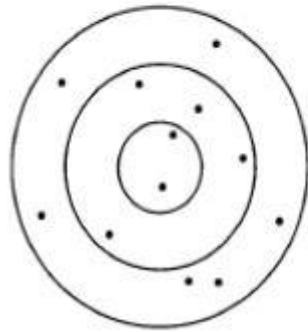
Random: small
Systematic: small

(a)

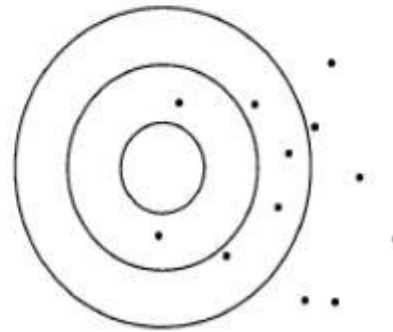


Random: small
Systematic: large

(b)



Random: large
Systematic: small



Random: large
Systematic: large

Errores aleatorios y sistematicos



¡Ésta es en general la situación en ciencia!

Tarea 1: Organización de la toma de datos (sitio y balanza de cada grupo)

1. Sodexo (Hall Universitario) : NT 2
2. Sodexo (La Pergola) : NT 1
3. Agronomía & Ingeniería Forestal : Terraillon 1
4. Economía y Administración : Cotidiana Azul
5. “Food Garden” : Profe (vieja)
6. Clementina : Cotidiana Blanca
7. Centro de Innovación Angelini : GAMA
8. Comedor de Ciencias Sociales : Terraillon 2