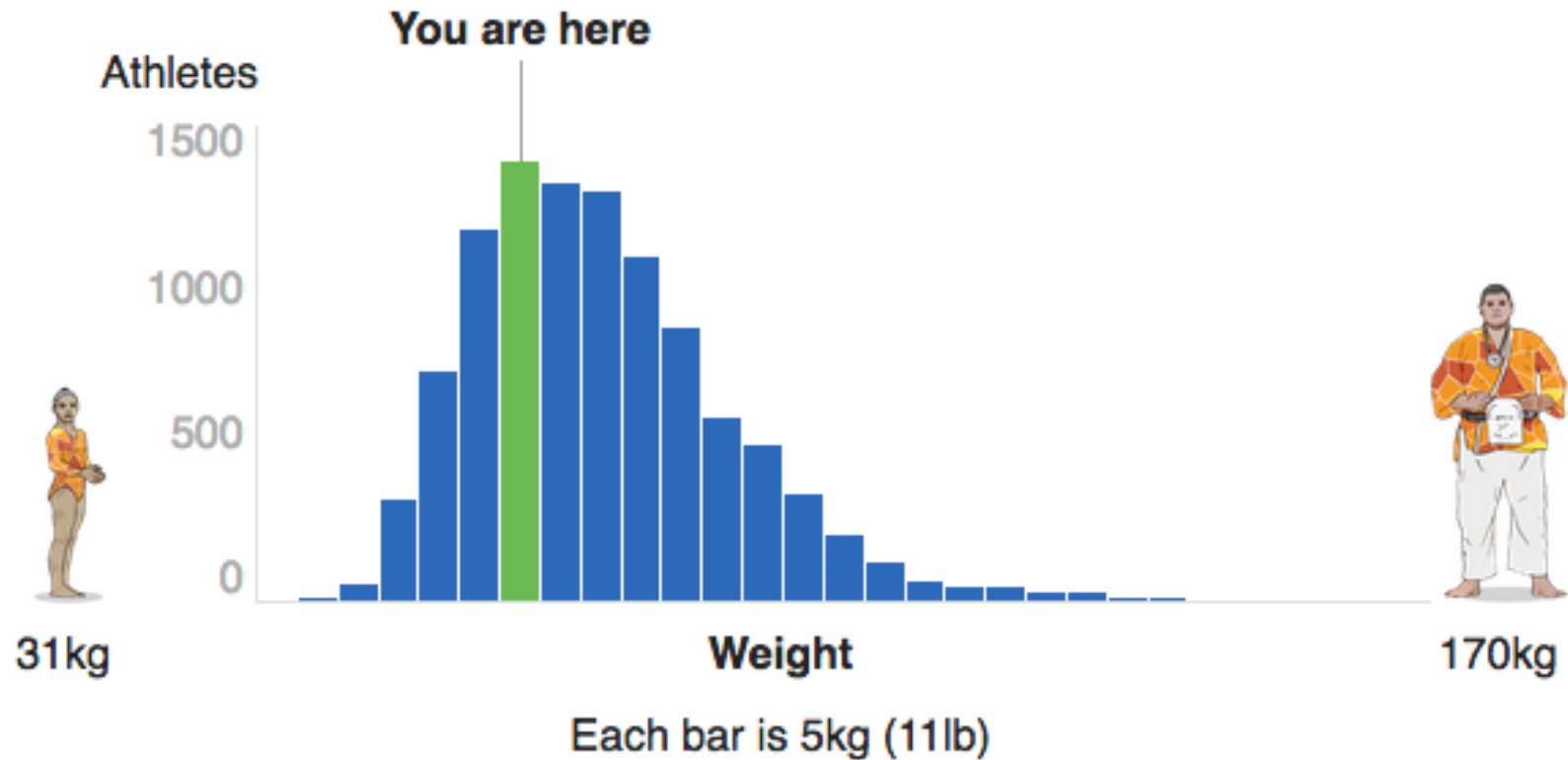


Histogramas

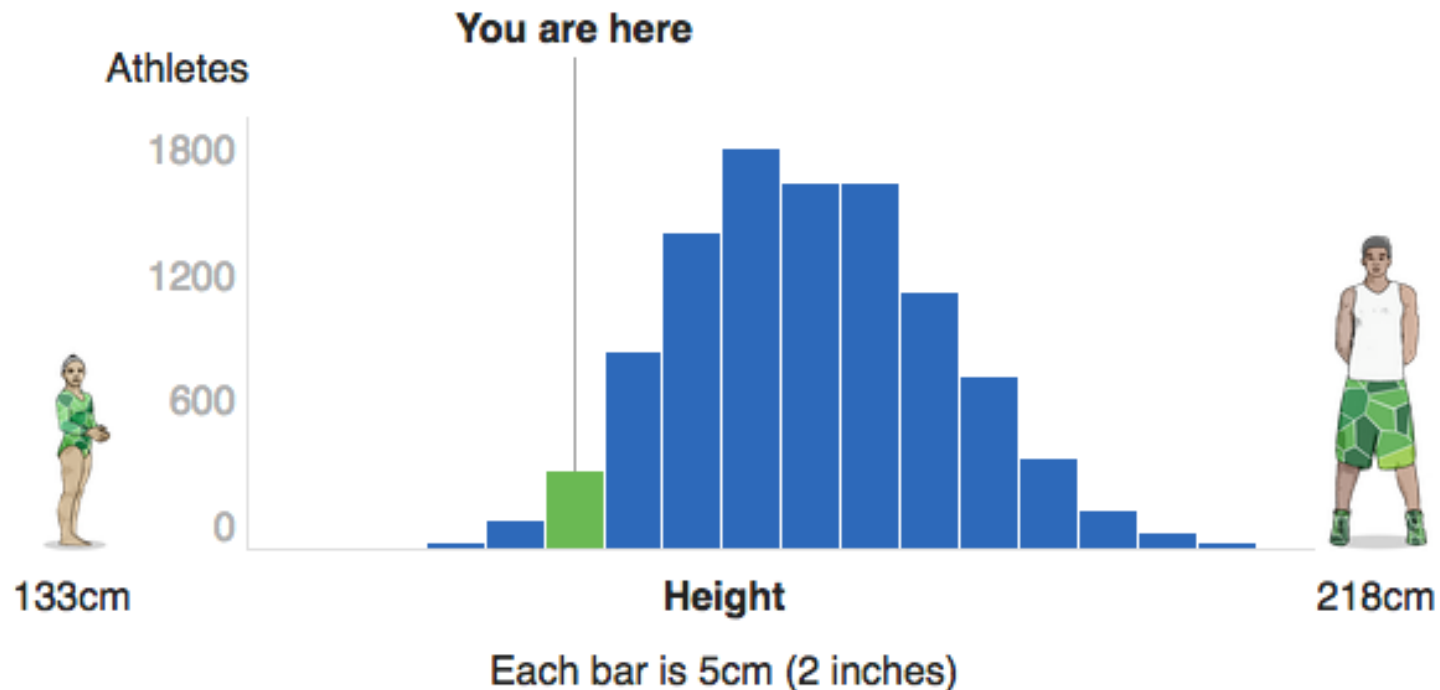
Elizabeth Hernández Marín
Laboratorio de Física
Febrero 2019

You are the same weight as 428 athletes. Flavia Saraiva is also among the lightest athletes at Rio, weighing in at a tiny 31kg (5st 1lb). Romanian judo contestant Daniel Natea is more than five times heavier at 170kg (26st 11lb). The average athlete weighs 72kg (11st 5lb).



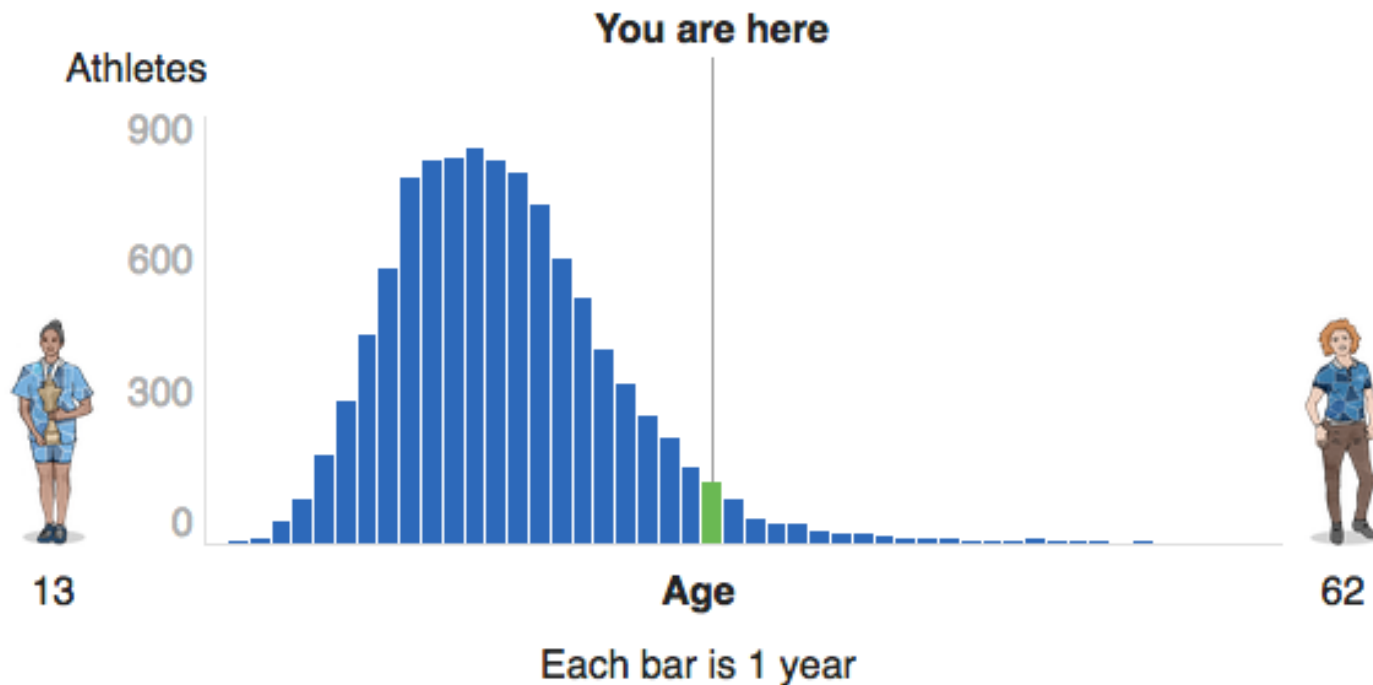
First your height...

You are the same height as 50 athletes. The shortest athlete in the database is Brazilian gymnast Flavia Saraiva at 1.33m (4ft 4in). The tallest athlete is Li Muhao, a Chinese basketball player who stands almost a metre taller at 2.18m (7ft 2in). The average height of athletes at the Games is 1.77m (5ft 10in).



Now your age

You are the same age as 141 athletes. The youngest competitor at the Games, Nepalese swimmer Gaurika Singh, is just 13. The oldest is equestrian Julie Brougham from New Zealand, who is taking part in her first Olympics at the age of 62. The average age of participants is 26. You share a birthday with 12 athletes.



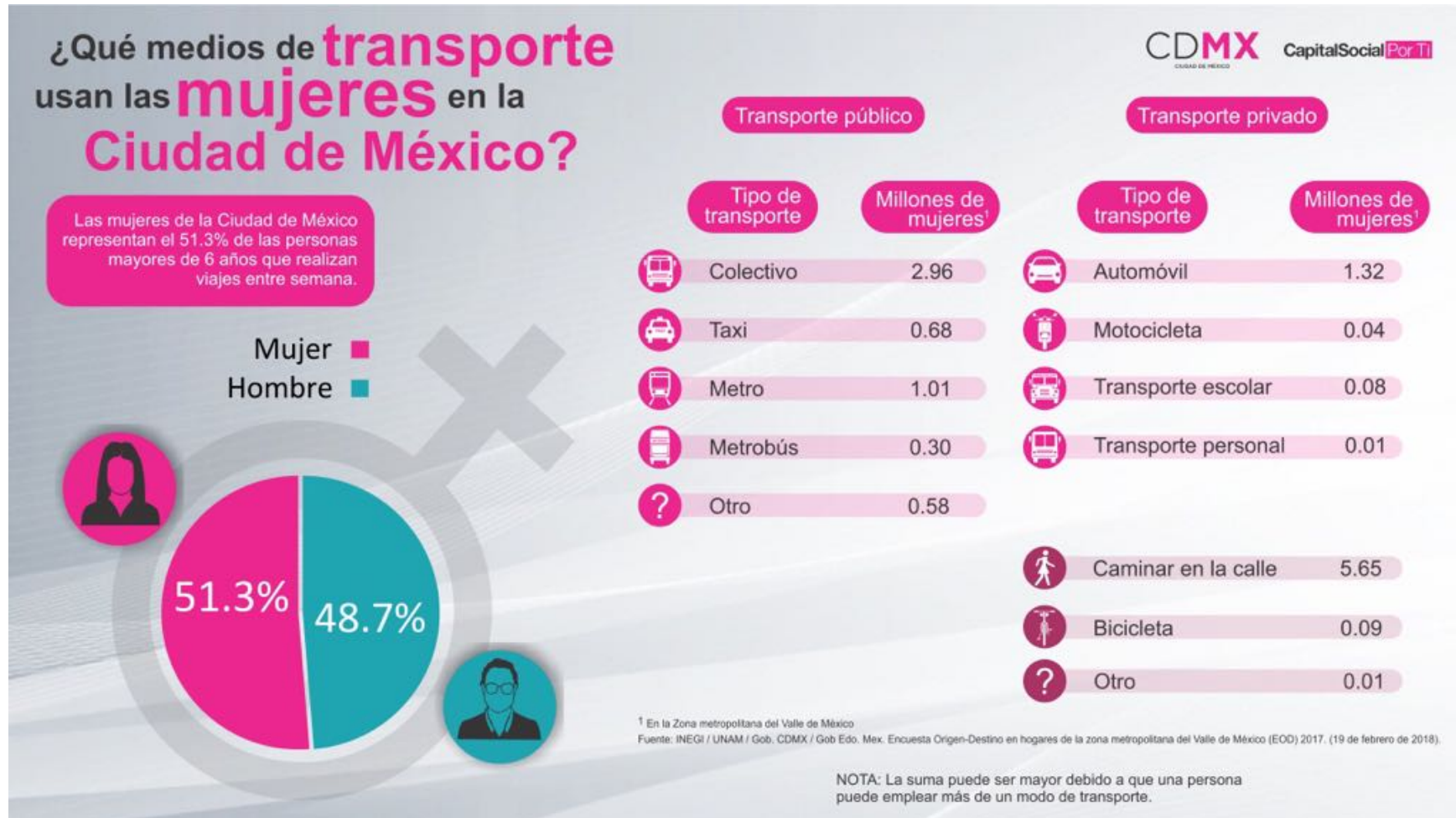
Frecuencias y su distribución

La **frecuencia** de una observación en particular, indica el número de veces que dicha observación se repite u ocurre en el conjunto de datos analizados.

La **distribución** de una variable es el patrón que siguen las frecuencias de la observación. Las distribuciones de frecuencia son representadas como *tablas de frecuencias, histogramas o polígonos*.

Tablas de frecuencias

Tabla que presenta datos estadísticos donde se anotan los valores de cierta característica (observación) junto con el número de veces (frecuencia) que dicho valor aparece en el conjunto de datos analizados.

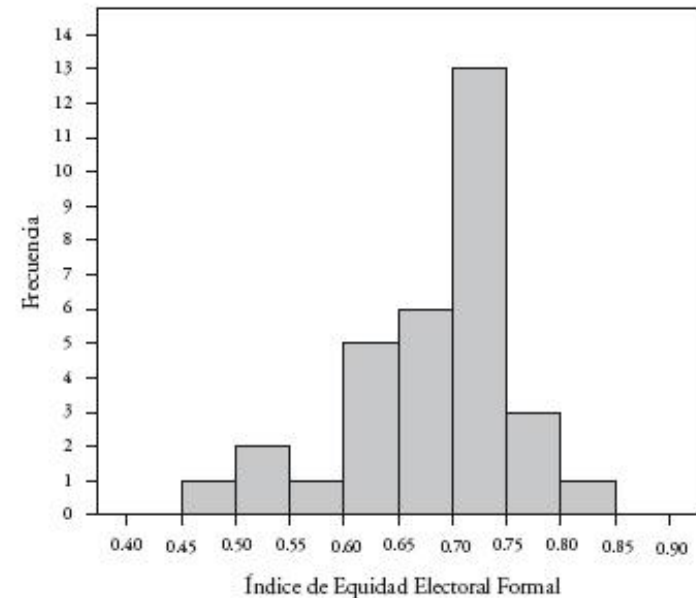


Histograma

Un histograma es una gráfica de la distribución de la frecuencia de una variable continua, basada en intervalos de clase.

Consiste en una serie de columnas, donde la altura de cada columna corresponde a la frecuencia de la ocurrencia del intervalo de clase que es la base que identifica a cada columna.

Gráfico 2
Histograma de los valores del índice EEF (2007)



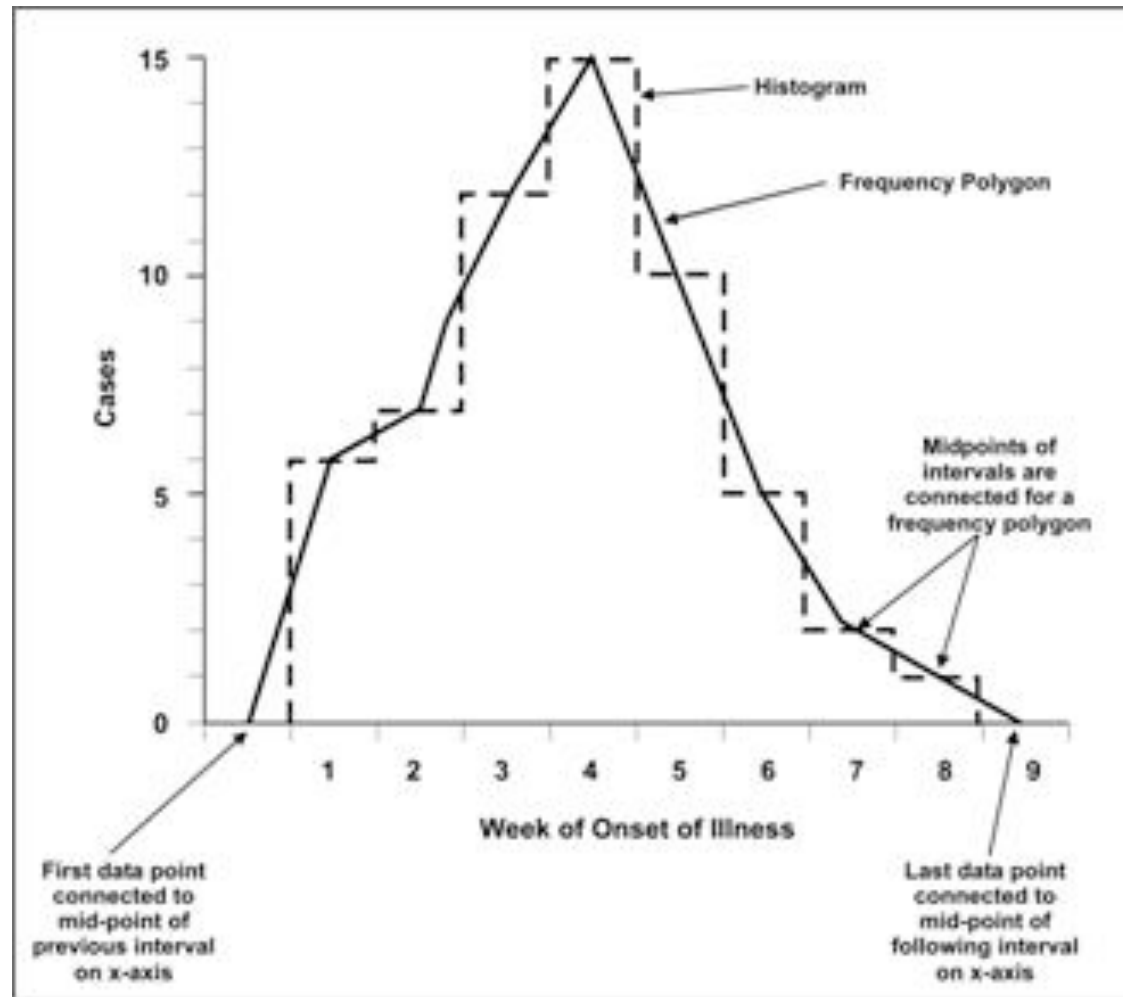
Fuente: elaboración propia a partir del análisis de la legislación electoral de las entidades federativas.

Notar que no hay espacios entre columnas adyacentes.

Figura tomada de: "La equidad electoral formal en las entidades federativas: México (1996-2007)" Alejandro Monsiváis Carrillo

Polígonos de frecuencia

Gráfica que se forma al unir los puntos medios de la parte superior de las columnas del histograma.



Tamaño de la muestra (n): Cantidad total de los datos analizados.

*Para calcular la mediana se deben primero ordenar los datos en orden ascendente

valor que corresponde al ordinal resultante de: $(n+1)/2$

Ejemplos:

(a) Tiempo en que un atleta corre 200 m:

26.1 s, 25.6 s, 25.7 s, 25.2 s, 25.0 s.

1. Ordenar: (25.0, 25.2, 25.6, 25.7, 26.1)s

$(n + 1)/2$

$= (5 + 1) \div 2$

$= 3$

Mediana: valor que corresponde al **3er** dato = 25.6 s.

*Para calcular la mediana se deben primero ordenar los datos en orden ascendente

valor que corresponde al ordinal resultante de: $(n+1)/2$

Ejemplos:

(b) Tiempo en que un atleta corre 200 m:

26.1 s, 25.6 s, 25.7 s, 25.2 s, 25.0 s, 24.7.

1. Ordenar: (24.7, 25.0, 25.2, 25.6, 25.7, 26.1)s

$(n + 1)/2$

$= (6 + 1) \div 2$

$= 3.5$

Mediana: valor que corresponde al **promedio del 3er y 4to** datos

$= (25.2 + 25.6)s/2$

$= 25.4 \text{ s}$

Ejemplo. **valor que corresponde al ordinal**
resultante de: **$(137+1)/2 = 69$**

Ejercicio 1. Considere el siguiente conjunto de mediciones de longitud. La unidad está en cm.

0.06	1.04	1.21	1.88	2.06	2.10	2.14	2.27	2.47	2.60
2.70	2.89	3.07	3.09	3.30	3.31	3.50	3.71	3.78	3.86
3.93	4.06	4.15	4.19	4.22	4.38	4.40	4.45	4.55	4.57
4.66	4.77	4.86	4.86	4.87	4.91	4.98	5.05	5.08	5.08
5.13	5.19	5.19	5.22	5.37	5.40	5.43	5.49	5.55	5.57
5.59	5.59	5.66	5.68	5.69	5.71	5.72	5.75	5.90	5.91
5.91	6.03	6.13	6.21	6.29	6.29	6.42	6.44	6.50	6.57
6.60	6.68	6.72	6.78	6.83	6.84	6.85	6.86	7.03	7.11
7.26	7.31	7.61	7.63	7.64	7.67	7.72	7.78	7.83	7.85
7.86	7.88	7.89	7.97	7.99	8.04	8.10	8.19	8.23	8.23
8.28	8.39	8.39	8.51	8.57	8.57	8.62	8.62	8.66	8.66
8.69	8.76	9.10	9.19	9.24	9.28	9.35	9.50	9.60	9.69
9.76	9.76	9.86	10.11	10.12	10.58	10.65	10.83	11.13	11.21
11.30	11.43	11.48	11.99	12.01	12.44	13.59			

Mediana: 6.50 cm

Moda: Dato (o datos) que más se repiten.

Ejercicio 1. Considere el siguiente conjunto de mediciones de longitud. La unidad está en cm.

0.06	1.04	1.21	1.88	2.06	2.10	2.14	2.27	2.47	2.60
2.70	2.89	3.07	3.09	3.30	3.31	3.50	3.71	3.78	3.86
3.93	4.06	4.15	4.19	4.22	4.38	4.40	4.45	4.55	4.57
4.66	4.77	4.86	4.86	4.87	4.91	4.98	5.05	5.08	5.08
5.13	5.19	5.19	5.22	5.37	5.40	5.43	5.49	5.55	5.57
5.59	5.59	5.66	5.68	5.69	5.71	5.72	5.75	5.90	5.91
5.91	6.03	6.13	6.21	6.29	6.29	6.42	6.44	6.50	6.57
6.60	6.68	6.72	6.78	6.83	6.84	6.85	6.86	7.03	7.11
7.26	7.31	7.61	7.63	7.64	7.67	7.72	7.78	7.83	7.85
7.86	7.88	7.89	7.97	7.99	8.04	8.10	8.19	8.23	8.23
8.28	8.39	8.39	8.51	8.57	8.57	8.62	8.62	8.66	8.66
8.69	8.76	9.10	9.19	9.24	9.28	9.35	9.50	9.60	9.69
9.76	9.76	9.86	10.11	10.12	10.58	10.65	10.83	11.13	11.21
11.30	11.43	11.48	11.99	12.01	12.44	13.59			

Es una distribución multimodal

"4.86" "5.08" "5.19" "5.59" "5.91" "6.29" "8.23" "8.39" "8.57" "8.62" "8.66" "9.76"

Tabla 2. Masa (g) de un lote de chicles marca Bubli-Bubli. Los datos fueron colectados en el semestre 2015-2. Incertidumbre de las medidas ± 0.01 g

4.46	5.92	6.11	6.30	6.53
4.94	5.95	6.16	6.30	6.54
5.37	5.98	6.18	6.31	6.60
5.44	6.00	6.18	6.32	6.60
5.52	6.00	6.20	6.32	6.65
5.65	6.00	6.20	6.33	6.69
5.66	6.00	6.20	6.35	6.70
5.70	6.00	6.22	6.35	6.75
5.70	6.00	6.24	6.35	6.76
5.73	6.02	50 6.24	6.36	6.76
5.79	6.03	51 6.25	6.38	6.78
5.80	6.04	6.25	6.40	6.83
5.81	6.08	6.26	6.45	6.87
5.83	6.08	6.27	6.46	6.88
5.85	6.10	6.28	6.49	6.92
5.85	6.10	6.28	6.49	6.93
5.85	6.10	6.28	6.49	6.93
5.90	6.10	6.28	6.49	6.96
5.91	6.10	6.28	6.50	6.99
5.92	6.10	6.28	6.52	7.02

valor que corresponde al ordinal resultante de:

$$(100+1)/2=50.5$$

Datos 50° y 51

Mediana: $(6.24+6.25)/2 = 6.245$ g

Moda: 6.00 g, 6.10 g y 6.28 g

Más definiciones

Ejercicio 1. Considere el siguiente conjunto de mediciones de longitud. La unidad está en cm.

0.06	1.04	1.21	1.88	2.06	2.10	2.14	2.27	2.47	2.60
2.70	2.89	3.07	3.09	3.30	3.31	3.50	3.71	3.78	3.86
3.93	4.06	4.15	4.19	4.22	4.38	4.40	4.45	4.55	4.57
4.66	4.77	4.86	4.86	4.87	4.91	4.98	5.05	5.08	5.08
5.13	5.19	5.19	5.22	5.37	5.40	5.43	5.49	5.55	5.57
5.59	5.59	5.66	5.68	5.69	5.71	5.72	5.75	5.90	5.91
5.91	6.03	6.13	6.21	6.29	6.29	6.42	6.44	6.50	6.57
6.60	6.68	6.72	6.78	6.83	6.84	6.85	6.86	7.03	7.11
7.26	7.31	7.61	7.63	7.64	7.67	7.72	7.78	7.83	7.85
7.86	7.88	7.89	7.97	7.99	8.04	8.10	8.19	8.23	8.23
8.28	8.39	8.39	8.51	8.57	8.57	8.62	8.62	8.66	8.66
8.69	8.76	9.10	9.19	9.24	9.28	9.35	9.50	9.60	9.69
9.76	9.76	9.86	10.11	10.12	10.58	10.65	10.83	11.13	11.21
11.30	11.43	11.48	11.99	12.01	12.44	13.59			

Intervalo principal = (medida de mayor valor) - (medida de menor valor)

Ejemplo: 13.59 cm - 0.06 cm = 13.53 cm

Si una variable abarca un gran número de valores, es más práctico presentar los datos en grupos de valores (**intervalos, o intervalos de clases**).

Todo intervalo debe quedar definido por un límite inferior y uno superior.

Para determinar el número de clases se puede emplear:

Regla de Sturges. $1 + 3.3 \log_{10}(n)$

Regla de Rice. $2\sqrt[3]{n}$

Para determinar el número de clases se puede emplear:

Regla de Sturges. $1 + 3.3 \log_{10}(n)$

Ejemplo: $1 + 3.3 \log_{10}(137) = 8.0511 \approx 8$

Regla de Rice. $2\sqrt[3]{n}$

Ejemplo: $2\sqrt[3]{137} = 10.31 \approx 10$

Tamaño de la muestra (n): 137

Intervalo principal: 13.59 cm - 0.06 cm = 13.53 cm

Regla de Sturges: $1+3.3\log(137) = 8.05 \approx 8$

Generar una tabla de frecuencias con 8 intervalos de clase

Tamaño del intervalo= $\frac{13.53 \text{ cm}}{8} = 1.69125 \text{ cm}$

ó Tamaño de clase

1.69*8=13.52 😞

1.70*8=13.6 😞 Se tendría que ampliar el
intervalo principal por ejemplo
de 0.03 a 13.63

1.692*8=13.536 😞

ampliar el intervalo principal de
0.057-13.593

Tabla de frecuencias

Generar una tabla de frecuencias con 8 intervalos de clase

Intervalo , cm	frecuencia
[0,057,1.749]	
(1.749,3.441]	
(3.441,5.133]	
(5.133,6.825]	
(6.825,8.517]	
(8.517,10.209]	
(10.209,11.901]	
(11.901,13.593]	

Tamaño del intervalo 1.692 cm

ampliar el intervalo principal de 0.057-13.593

$$0.057 + 1.692 = 1.749$$

$$1.749 + 1.692 = 3.441$$

Tabla de frecuencias

Generar una tabla de frecuencias con 8 intervalos de clase

Intervalo , cm	frecuencia
[0,057,1.749]	3
(1.749,3.441]	13
(3.441,5.133]	25
(5.133,6.825]	33
(6.825,8.517]	30
(8.517,10.209]	21
(10.209,11.901]	8
(11.901,13.593]	4

Ejercicio 1. Considere el siguiente conjunto de mediciones de longitud. La unidad está en cm.

0.06	1.04	1.21	1.88	2.06	2.10	2.14	2.27	2.47	2.60
2.70	2.89	3.07	3.09	3.30	3.31	3.50	3.71	3.78	3.86
3.93	4.06	4.15	4.19	4.22	4.38	4.40	4.45	4.55	4.57
4.66	4.77	4.86	4.86	4.87	4.91	4.98	5.05	5.08	5.08
5.13	5.19	5.19	5.22	5.37	5.40	5.43	5.49	5.55	5.57
5.59	5.59	5.66	5.68	5.69	5.71	5.72	5.75	5.90	5.91
5.91	6.03	6.13	6.21	6.29	6.29	6.42	6.44	6.50	6.57
6.60	6.68	6.72	6.78	6.83	6.84	6.85	6.86	7.03	7.11
7.26	7.31	7.61	7.63	7.64	7.67	7.72	7.78	7.83	7.85
7.86	7.88	7.89	7.97	7.99	8.04	8.10	8.19	8.23	8.23
8.28	8.39	8.39	8.51	8.57	8.57	8.62	8.62	8.66	8.66
8.69	8.76	9.10	9.19	9.24	9.28	9.35	9.50	9.60	9.69
9.76	9.76	9.86	10.11	10.12	10.58	10.65	10.83	11.13	11.21
11.30	11.43	11.48	11.99	12.01	12.44	13.59			

clase modal

Puede haber una o más clases modales

*La suma de las frecuencias debe ser igual al tamaño de la muestra

Ejemplo

Ejercicio 1. Considere el siguiente conjunto de mediciones de longitud. La unidad está en cm.

0.06	1.04	1.21	1.88	2.06	2.10	2.14	2.27	2.47	2.60
2.70	2.89	3.07	3.09	3.30	3.31	3.50	3.71	3.78	3.86
3.93	4.06	4.15	4.19	4.22	4.38	4.40	4.45	4.55	4.57
4.66	4.77	4.86	4.86	4.87	4.91	4.98	5.05	5.08	5.08
5.13	5.19	5.19	5.22	5.37	5.40	5.43	5.49	5.55	5.57
5.59	5.59	5.66	5.68	5.69	5.71	5.72	5.75	5.90	5.91
5.91	6.03	6.13	6.21	6.29	6.29	6.42	6.44	6.50	6.57
6.60	6.68	6.72	6.78	6.83	6.84	6.85	6.86	7.03	7.11
7.26	7.31	7.61	7.63	7.64	7.67	7.72	7.78	7.83	7.85
7.86	7.88	7.89	7.97	7.99	8.04	8.10	8.19	8.23	8.23
8.28	8.39	8.39	8.51	8.57	8.57	8.62	8.62	8.66	8.66
8.69	8.76	9.10	9.19	9.24	9.28	9.35	9.50	9.60	9.69
9.76	9.76	9.86	10.11	10.12	10.58	10.65	10.83	11.13	11.21
11.30	11.43	11.48	11.99	12.01	12.44	13.59			

media (promedio) $\mu = 6.624015$ cm

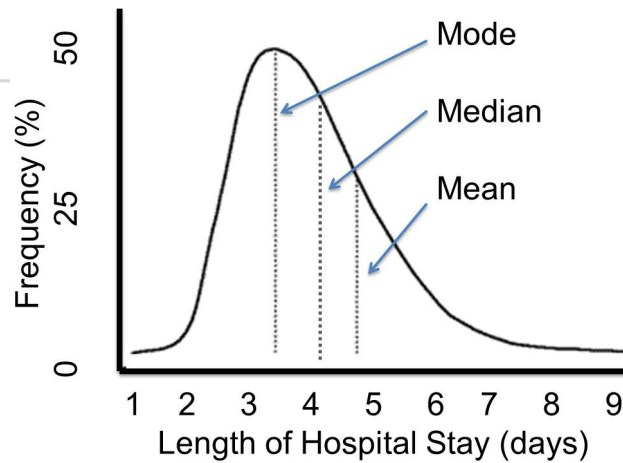
desviación estándar $s = 2.645935$ cm

$\mu - \sigma = 3.97809$ cm

$\mu + \sigma = 9.26995$

mediana = 6.50 cm

Distribución de frecuencia



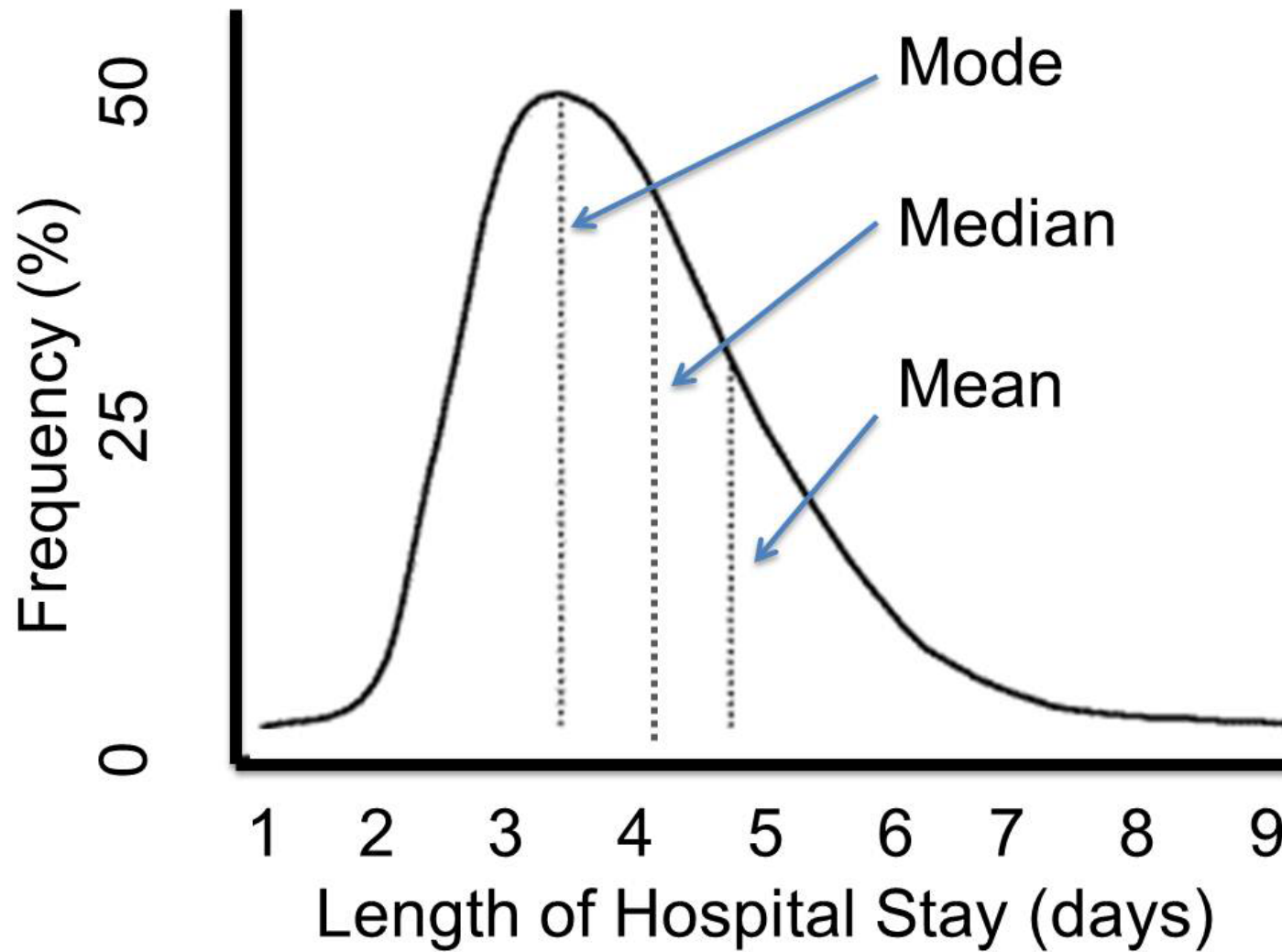
%frecuencia = 100(frecuencia de una clase)/(tamaño de la muestra) $\%f_i = \frac{100 \cdot f_i}{n}$

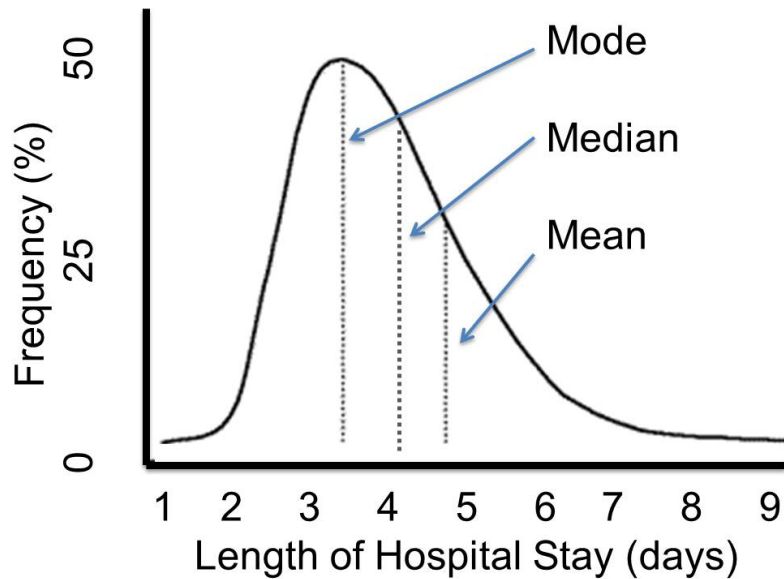
media (promedio) $\bar{x} = \frac{\sum x_i}{n}$

moda (o clase modal): Corresponde al valor (o intervalo) que tiene la mayor frecuencia

mediana: Corresponde al intervalo que se encuentra a la mitad de los datos

Distribución de frecuencia





¿Existe alguna distribución cuya media, mediana y moda sean las mismas?

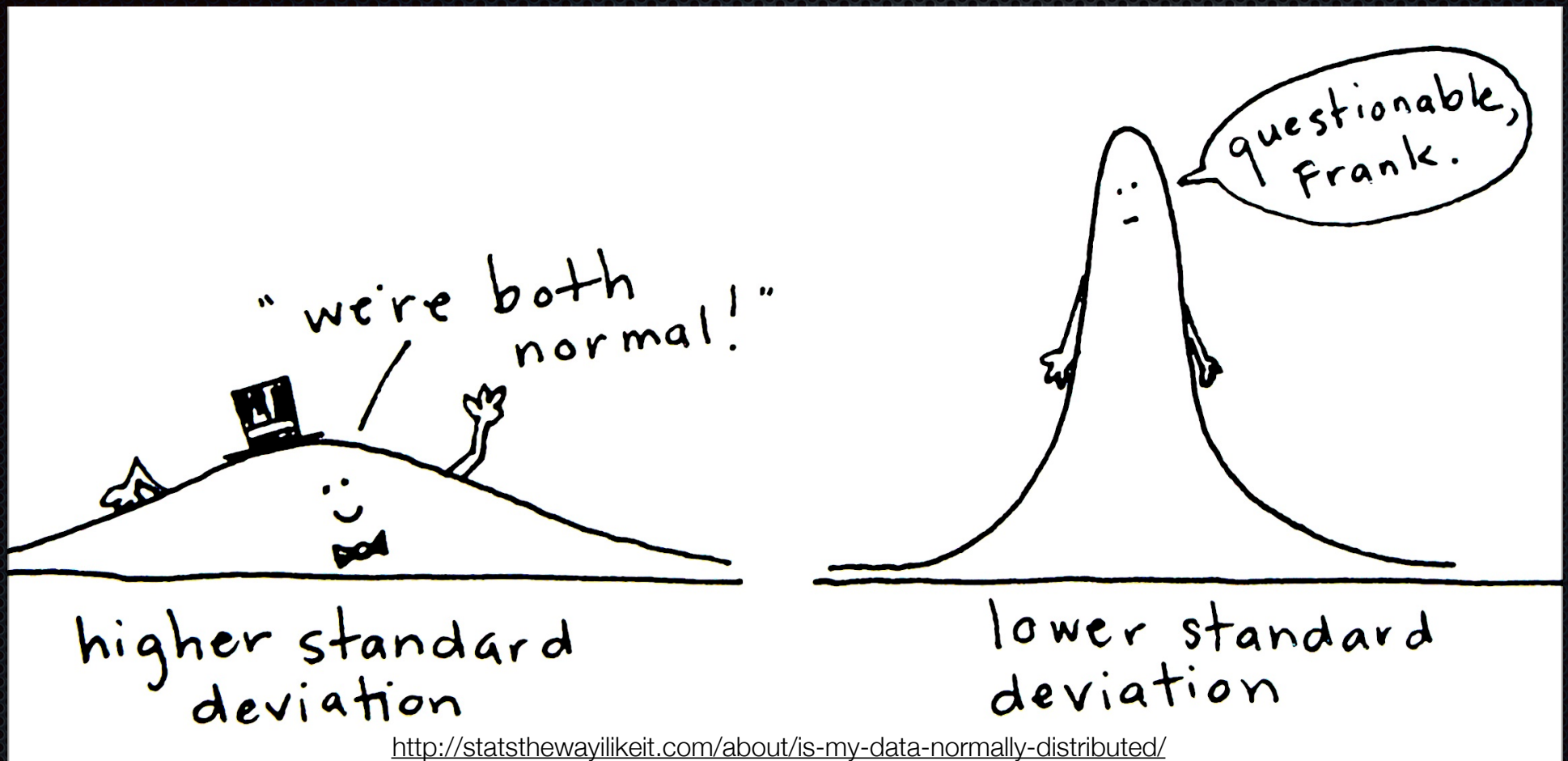


NORMAL DISTRIBUTION



PARANORMAL DISTRIBUTION

F. Freeman



Distribución gaussiana

Elizabeth Hernández Marín
Laboratorio de Física

Función gaussiana

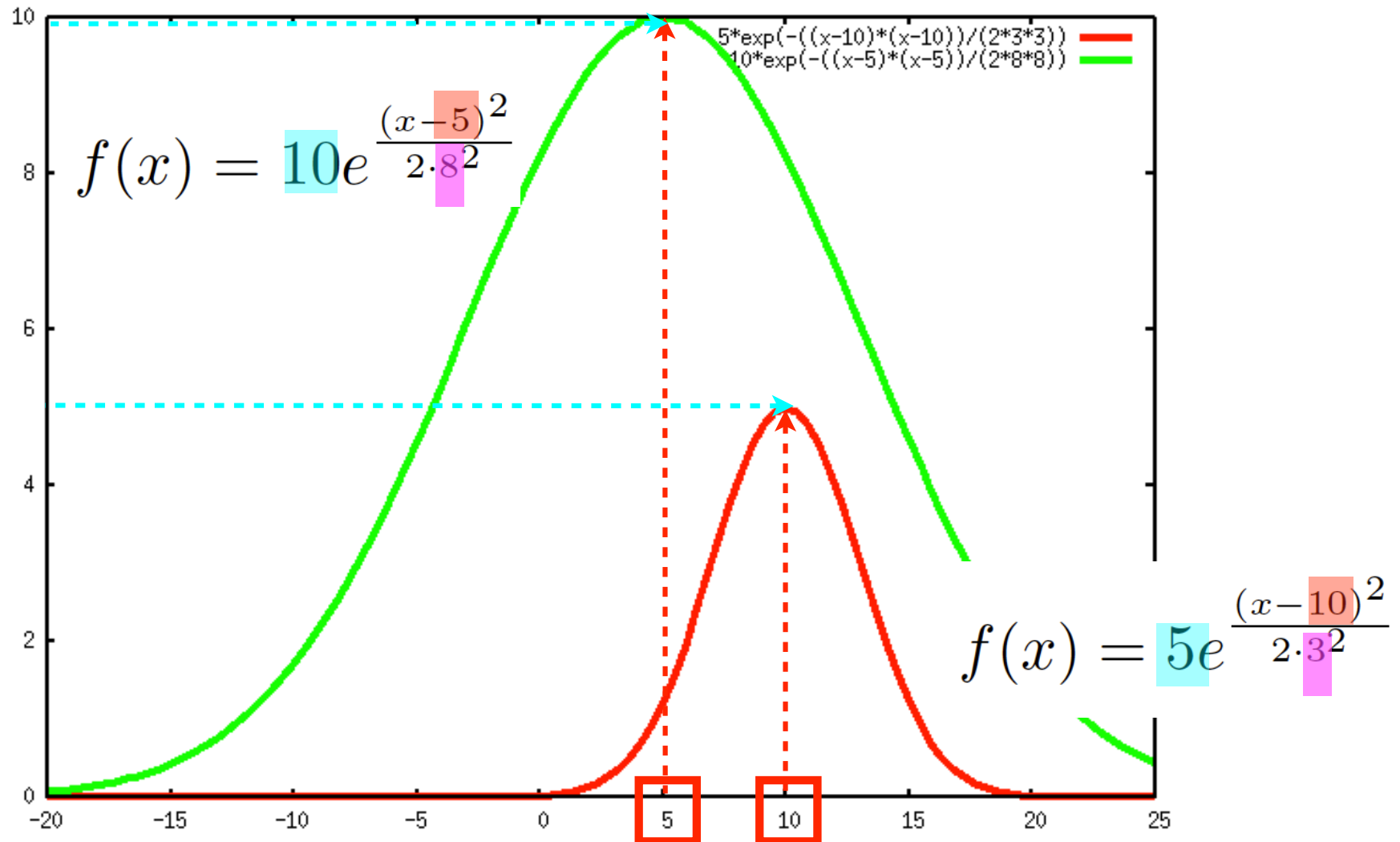
$$f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$$

Máximo: $f'(x) = 0 \implies x = b$ $f(b) = a$

$f''(x) = 0$ Puntos de inflexión: $b - c$
 $b + c$

Función gaussiana

$$f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$$



Ejemplo

Ejercicio 1. Considere el siguiente conjunto de mediciones de longitud. La unidad está en cm.

0.06	1.04	1.21	1.88	2.06	2.10	2.14	2.27	2.47	2.60
2.70	2.89	3.07	3.09	3.30	3.31	3.50	3.71	3.78	3.86
3.93	4.06	4.15	4.19	4.22	4.38	4.40	4.45	4.55	4.57
4.66	4.77	4.86	4.86	4.87	4.91	4.98	5.05	5.08	5.08
5.13	5.19	5.19	5.22	5.37	5.40	5.43	5.49	5.55	5.57
5.59	5.59	5.66	5.68	5.69	5.71	5.72	5.75	5.90	5.91
5.91	6.03	6.13	6.21	6.29	6.29	6.42	6.44	6.50	6.57
6.60	6.68	6.72	6.78	6.83	6.84	6.85	6.86	7.03	7.11
7.26	7.31	7.61	7.63	7.64	7.67	7.72	7.78	7.83	7.85
7.86	7.88	7.89	7.97	7.99	8.04	8.10	8.19	8.23	8.23
8.28	8.39	8.39	8.51	8.57	8.57	8.62	8.62	8.66	8.66
8.69	8.76	9.10	9.19	9.24	9.28	9.35	9.50	9.60	9.69
9.76	9.76	9.86	10.11	10.12	10.58	10.65	10.83	11.13	11.21
11.30	11.43	11.48	11.99	12.01	12.44	13.59			

Calcular la media, el desvío estándar y la mediana de la muestra.

Tamaño de la muestra: 137

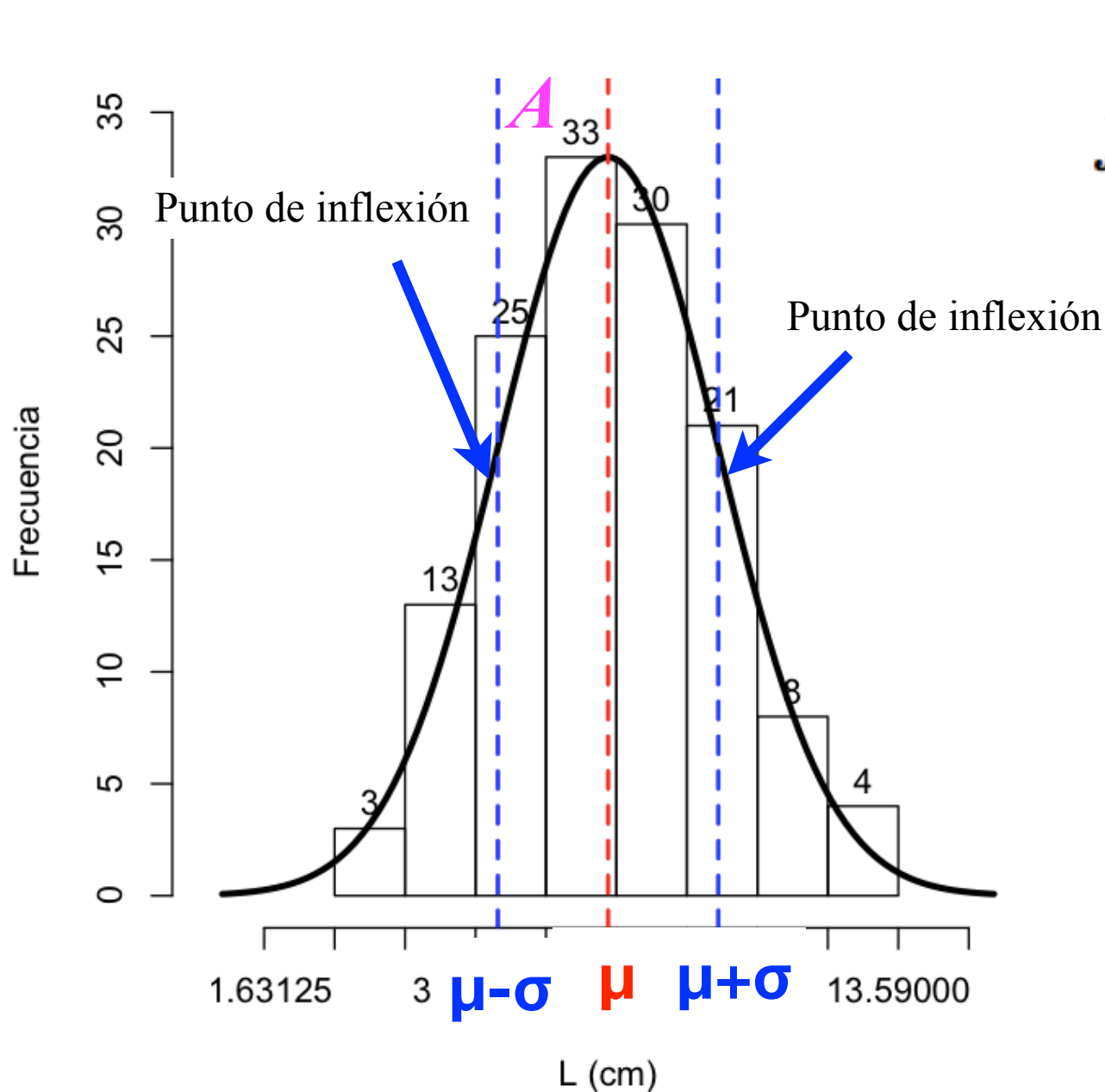
media: **6.624015** cm

mediana: **6.50 cm** (69º dato)

desv. estándar: **2.645934** cm

2.65 cm

distribución gaussiana



$$f(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$A = 33$$

Media

$$\mu = 6.62 \text{ cm}$$

Desviación estándar
 $\sigma = 2.65 \text{ cm}$

$$\mu - \sigma = 3.97 \text{ cm}$$

$$\mu + \sigma = 9.27 \text{ cm}$$

$$f(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Distribución NORMAL

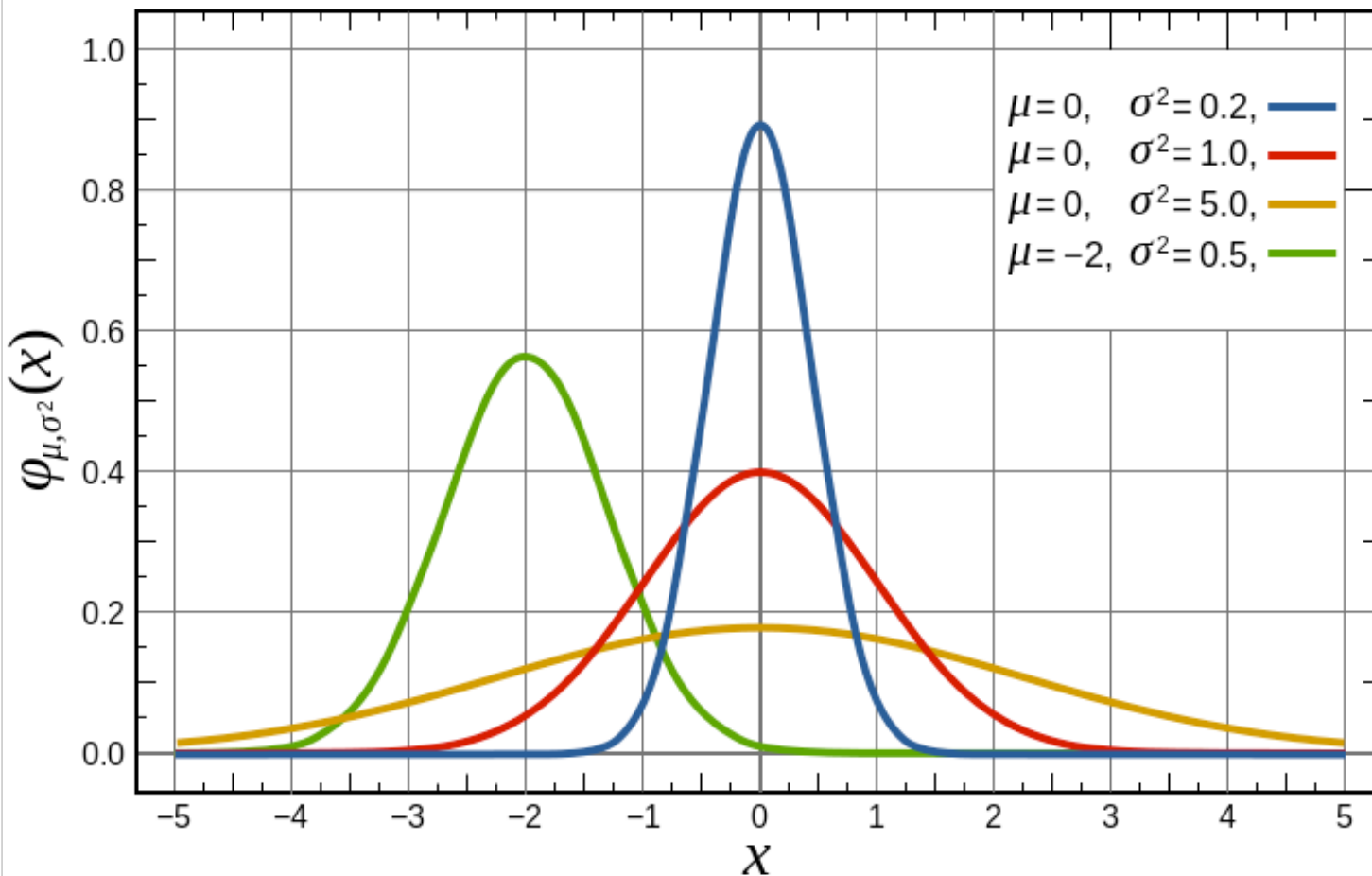
$$\int_{-\infty}^{\infty} e^{-a(x+b)^2} dx = \sqrt{\frac{\pi}{a}}$$

$$\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\sigma^2\pi} = \sigma\sqrt{2\pi}$$

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

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1$$

Distribución normal



Integral
bajo la
curva es
1

https://en.wikipedia.org/wiki/Gaussian_function#/media/File:Normal_Distribution_PDF.svg

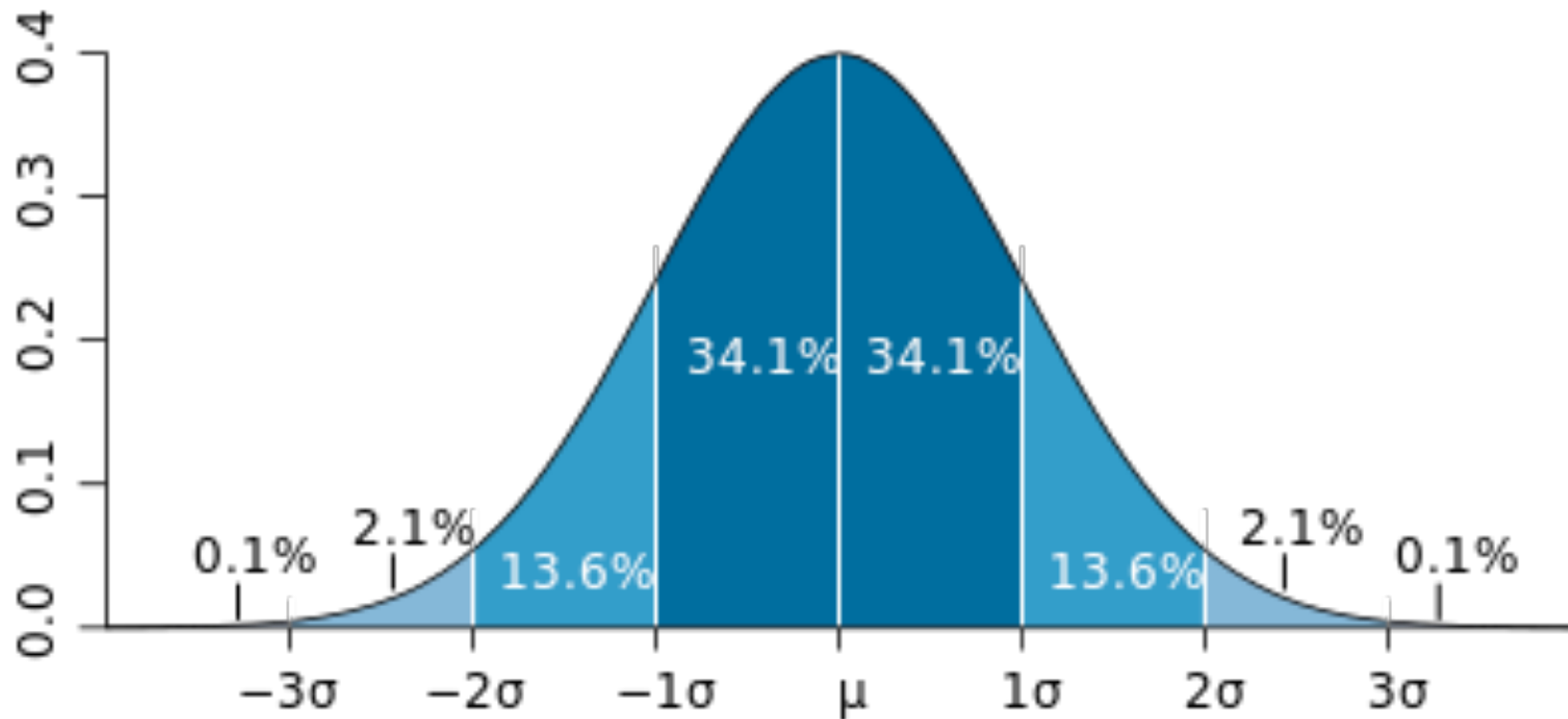
$$\varphi(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

μ , media
 σ , desv. est.

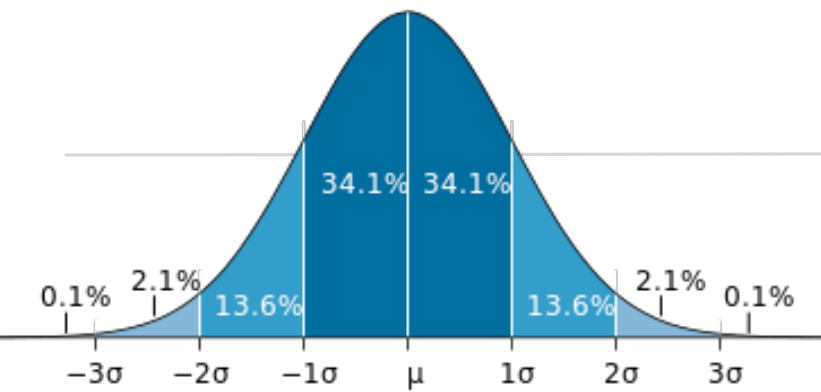
Distribución NORMAL

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

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1$$



Distribución gaussiana



68.2% de los datos se encuentran en el intervalo $[\mu - \sigma, \mu + \sigma]$

95.4% de los valores se encuentran en el intervalo $[\mu - 2\sigma, \mu + 2\sigma]$

99.6% de los valores se encuentran en el intervalo $[\mu - 3\sigma, \mu + 3\sigma]$

Actividad individual en clase

Realizar la tabla de frecuencias y el histograma a partir de los resultados obtenidos con la medición de su lote de dulces

Tarea 1

Realizar el histograma con el programa R. El código se proporciona en la página de amygd. Imprimir el histograma que obtuvo.

Tarea 2. Ejercicios de incertidumbres (medidas directas e indirectas) en google classroom