CONOCIENDO LA DISTRIBUCIÓN NORMAL

Para conocer profundamente el concepto de distribución de densidad de probabilidad y más concretamente de distribución normal solo necesitas tocar los datos y ver qué está pasando.

Básicamente jugar un poco.

¡Vamos a ello!

La cuestión es que vas a crear tú mismo/a algunas distribuciones normales y las compararás entre sí.

Crea dos distribuciones normales (misma dispersión)

Crea una distribución normal “m1” con estas dos características:

- Valor central (media): 65 kg
- Dispersión (desviación estándar): 12 kg

Y otra distribución normal “m2”, con estos dos parámetros:

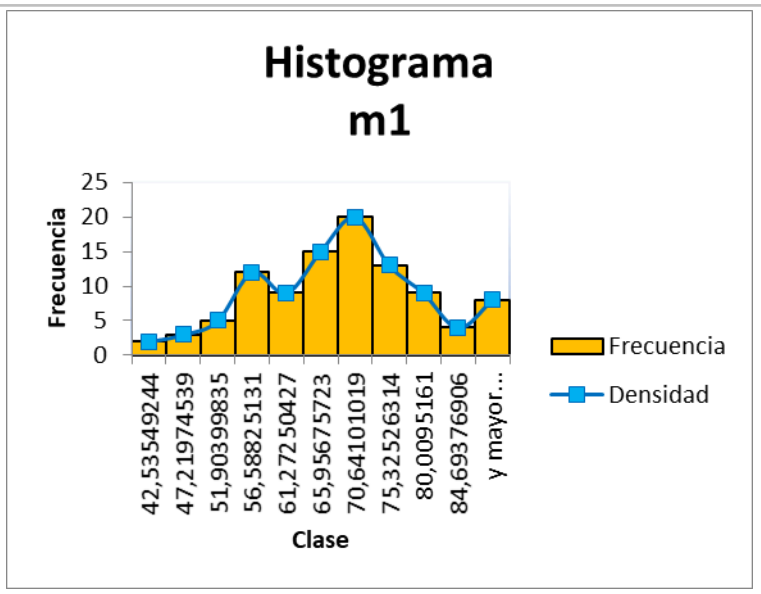
- Valor central (media): 42kg
- Dispersión (desviación estándar): 12 kg

Para crear estas muestras puedes ver el vídeo Tech Zen para calcular muestras de distribuciones normales.

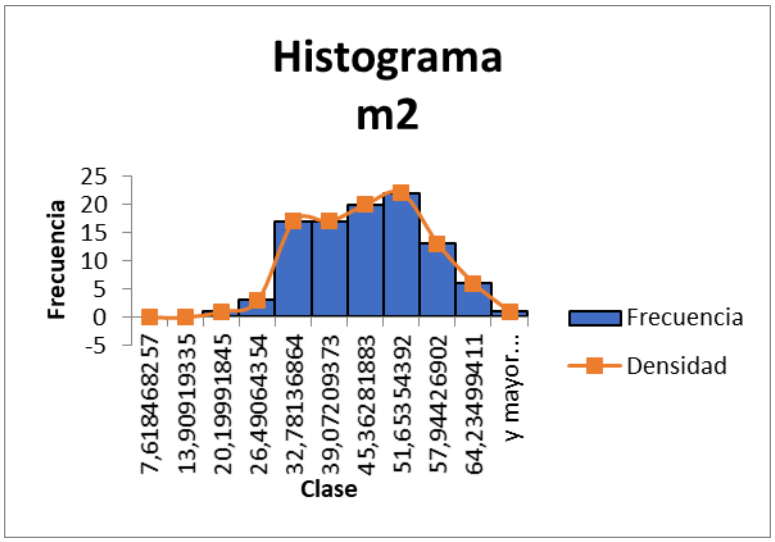
Dibuja los dos histogramas, calcula la media y la desviación estándar de las dos muestras de 100 observaciones y compara las dos distribuciones normales gráficamente y numéricamente.

Copias los dos histogramas y copia las medias y las desviaciones estándar que has calculado.

- Muestra m1
- Muestra m2



Con una media de: **64,6757178**
Con una desviación típica: **11,683423**



Con una media de: **41,0814314**
Con una desviación típica: **12,9805345**

¿Qué observas?

Podemos observar que los valores de ambas distribuciones se concentran alrededor de distintas medias por lo que su simetría es distinta sin embargo al tener, prácticamente, la misma desviación típica los datos se dispersan de la misma forma mostrando la misma amplitud en la distribución de sus valores. Son muestras cuya distribución de los valores con respecto de sus medias son muy parecidas.

Crea dos distribuciones normales (misma centralidad)

Crea otra distribución normal "m3", con estos dos parámetros:

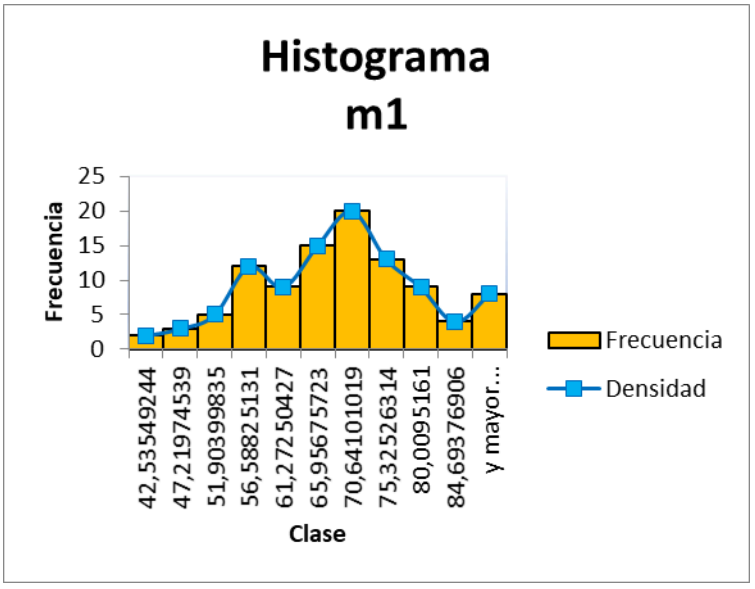
- Valor central (media): 65kg
- Dispersión (desviación estándar): 2 kg

Para crear estas muestras puedes ver el vídeo Tech Zen para calcular muestras de distribuciones normales.

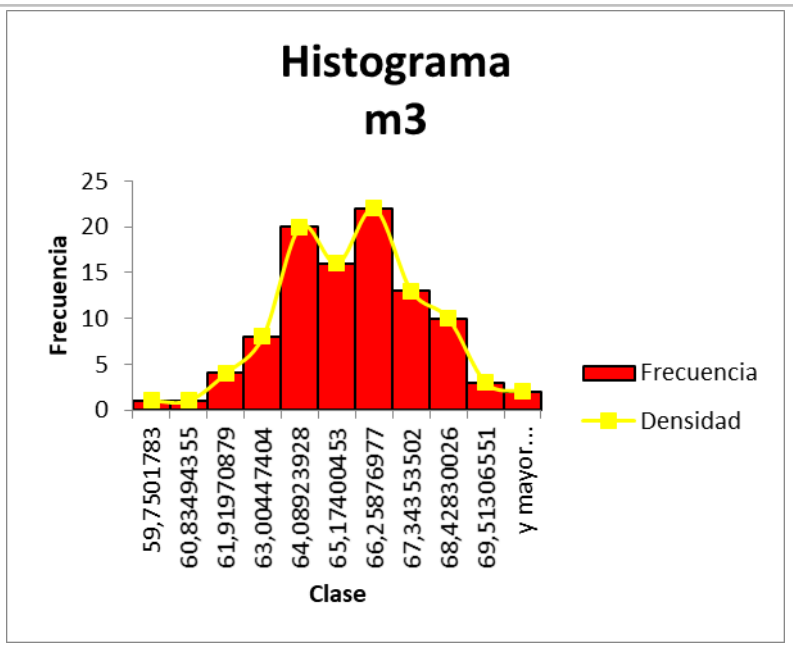
Dibuja los histogramas de la muestra "m1" y "m3", calcula la media y la desviación estándar de las dos muestras de 100 observaciones y compara las dos distribuciones normales.

Copias los dos histogramas y copia las medias y las desviaciones estándar que has calculado.

- Muestra m1
- Muestra m3



Con una media de: **64,7139287**
Con una desviación típica de: **12,1764376**



Con una media de: **65,0767221**
Con una desviación típica de: **2,15215735**

¿Qué observas?

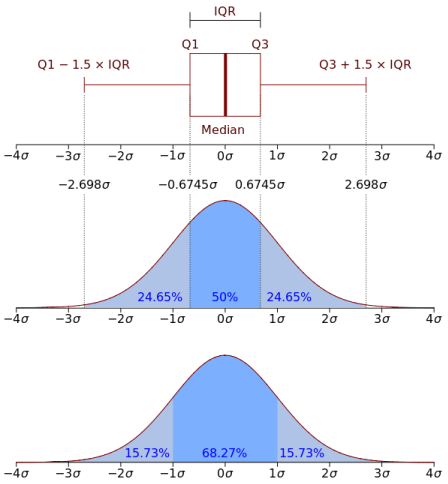
En este caso al tener las dos muestras la misma media podemos ver que ambos histogramas se parecen mucho ya que los datos están distribuidos alrededor de una misma media. Sin embargo, la forma de ambas distribuciones son distintas ya que sus desviaciones típicas son diferentes. Esto quiere decir que la dispersión de los datos con respecto a sus medias es distinta estableciendo curtosis diferentes. La muestra m1 es más mesocúrtica ya que sus datos se esparcen en mayor distancia respecto de su media mientras que la muestra m3 es más leptocúrtica debido a que sus datos se desvían menos de su media.

Aplicando las reglas de la distribución normal

De las tres muestras que has creado calcula las características de posición con la ayuda de la tabla.

Recuerda que esta distribución es especial. A la derecha tienes un esquema para recordarte algunas características

Calcula los resultados siguientes:



Característica		Muestra "m1"	Muestra "m2"	Muestra "m3"
CENTRALIDAD	Media	65,83247672	39,9278439	65,0767221
	Mediana (cuartil 2)	63,36004099	40,4326427	65,2065417
DISPERSIÓN	Desviación Estándar	12,50289066	11,5863404	2,15215735
POSICIÓN del 50% de los datos más centrados	=Media-0.67*Desv. Estándar	57,45554	32,1649958	63,6347767
	Cuartil 3	73,03529404	47,6594589	66,4705017
	=Media+0.67*Desv. Estándar	74,20941346	47,690692	66,5186675
	Cuartil 1	56,56596566	32,9572824	63,4401126
POSICIÓN del 95% de los datos más centrados	=Media+1.96*Desv. Estándar	90,33814242	62,6370711	69,2949505
	Percentil 97.5	87,91472336	66,0617306	69,4314029
	=Media-1.96*Desv. Estándar	41,32681101	17,2186167	60,8584937
	Percentil 2.5	44,98527488	16,9735958	61,2220968
FORMA	Curtosis	-0,64671291	0,48714925	-0,15627074
	Asimetría	0,309378238	-0,1899853	0,05418907

¿Qué patrón se repite en esta tabla?

Pista: media y mediana deberían ser parecidas, casi iguales

Para crear esta tabla utiliza las fórmulas que te enseñe en hojas de trabajo anteriores ☺

Aplicando la regla del 2

En la práctica normalmente se utiliza la regla del 2. (en lugar del 1.96 que has visto en la tabla anterior.

Si una variable cuantitativa presenta un histograma parecido a una distribución normal (con forma de montaña y simétrico) se aplica:

El 95% de los datos más centrados están entre la media +/- 2 veces la desviación estándar.

Es una regla efectiva para entender los límites de la variable rápidamente:

Característica		Muestra "m1"	Muestra "m2"	Muestra "m3"
POSICIÓN del 95% de los datos más centrados. Fórmula Exacta	=Media+1.96*Desv. Estándar	90,33814242	62,6370711	69,2949505
	Percentil 97.5	90,03048961	64,1635894	69,4314029
	=Media-1.96*Desv. Estándar	41,32681101	17,2186167	60,8584937
	Percentil 2.5	46,41684627	17,3646629	61,2220968
Regla del 2. POSICIÓN del 95% de los datos	=Media+2*Desv. Estándar	90,83825804	63,1005248	69,3810368
	=Media-2*Desv. Estándar	40,82669539	16,7551631	60,7724074

¡Buen trabajo!

¡Hoy has entendido con ejemplos la distribución normal y el concepto de distribución!