

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:

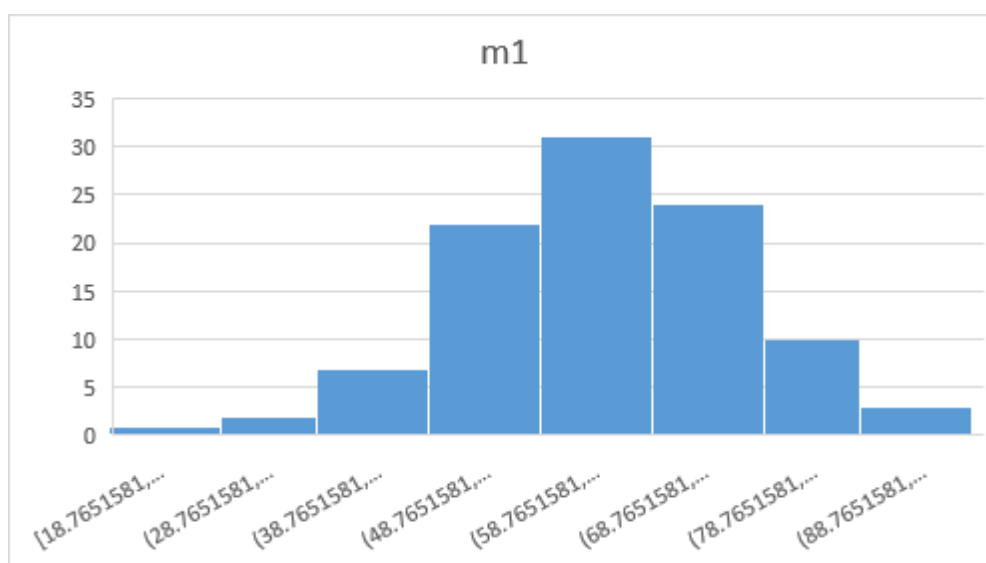
- Valor central (media): 42kg
- Dispersión (desviación estándar): 12 kg

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.

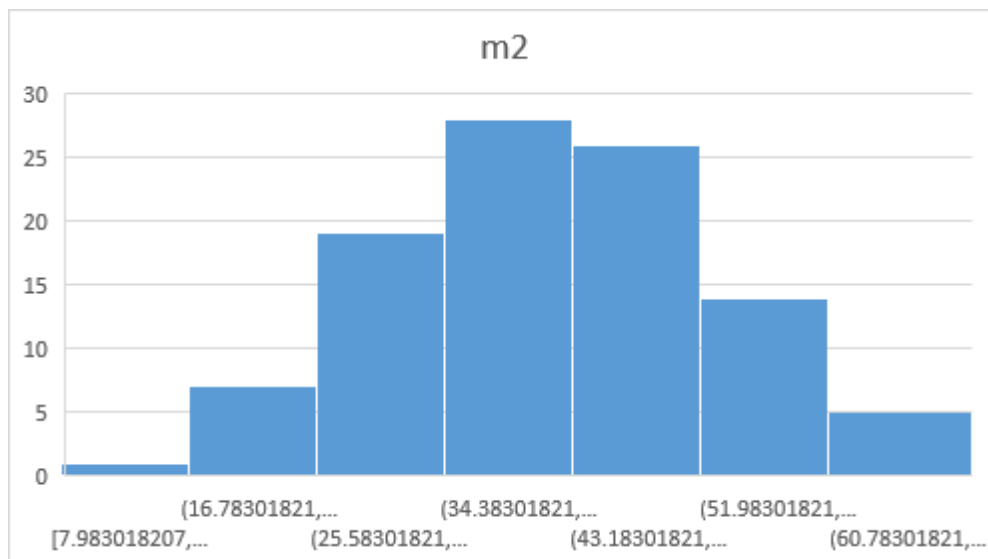
- Muestra m1



Media m1= 63.03

Desviación estándar m1= 13.16

- Muestra m2



Media m2= 41.22339209

Desviación estándar m2= 11.82565457

¿Qué observas?

Ambas muestras presentan un histograma de distribución de probabilidad con forma de campana bastante simétrica, siendo la forma de la m1 algo más aplanada y la m2 algo mas alargada.

Crea dos distribuciones normales (misma centralidad)

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

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

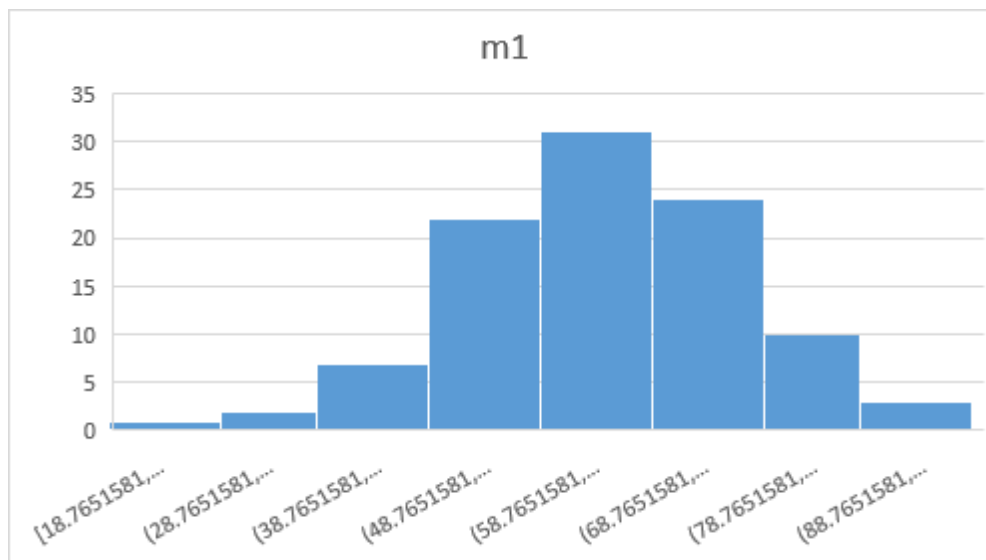
Para crear estas muestras puedes ver el vídeo Tech Zen para calcular muestras de distribuciones normales.

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.

CONOCIENDO LA DISTRIBUCIÓN NORMAL

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

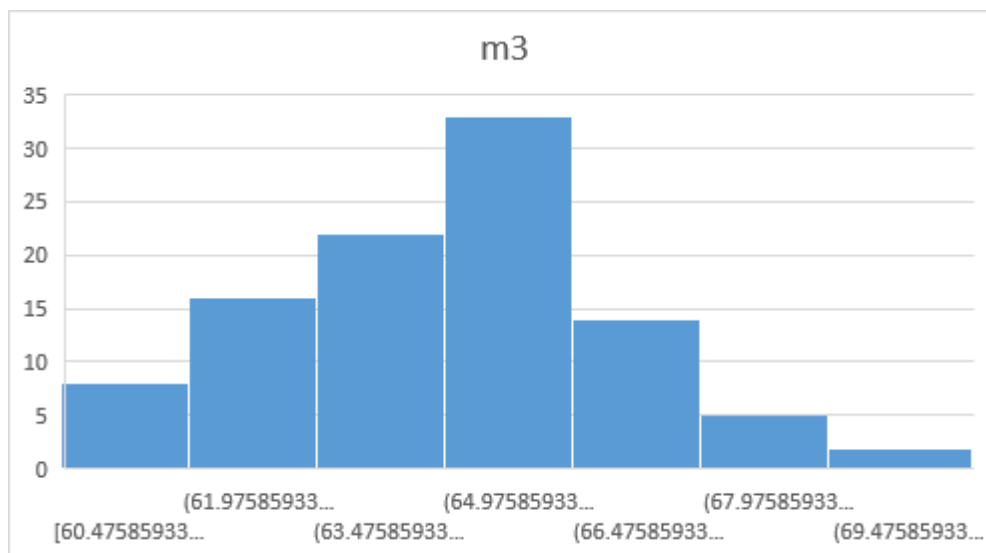
- Muestra m1



Media m1= 63.03

Desviación estándar m1= 13.16

- Muestra m3



Media m2= 64.86308879

Desviación estándar m2= 2.10961346

¿Qué observas?

CONOCIENDO LA DISTRIBUCIÓN NORMAL

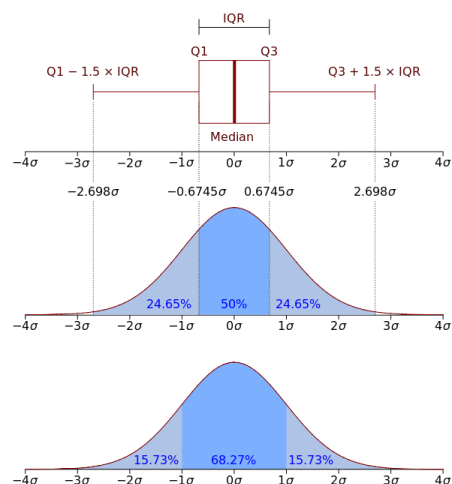
De nuevo, observamos dos curvas con forma de campana simétrica. En el caso de la m3, vemos que todos los valores están más cerca de la media, debido al valor bajo de la desviación estándar. Tiene menos dispersión que la m1.

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	63.0389447	41.2233921	64.8630888
	Mediana (cuartil 2)	62.6555313	41.0538581	64.8948622
DISPERSIÓN	Desviación Estándar	13.1681805	11.8256546	2.10961346
POSICIÓN del 50% de los datos más centrados	=Media-0.67*Desv. Estándar	54.2162637	33.3002035	63.4496478
	Cuartil 3	71.9070215	48.9074297	66.1538355
	=Media+0.67*Desv. Estándar	71.8616257	49.1465807	66.2765298
	Cuartil 1	53.811821	32.6557126	63.5282905
POSICIÓN del 95% de los datos más centrados	=Media+1.96*Desv. Estándar	88.8485786	64.4016751	68.9979312
	Percentil 97.5	93.3002249	64.0649241	69.0029716
	=Media-1.96*Desv. Estándar	37.2293108	18.0451091	60.7282464
	Percentil 2.5	38.3406551	17.2056814	60.0226049
FORMA	Curtosis	-0.1902233	-0.22721784	0.47346733
	Asimetría	0.19291884	-0.02210245	-0.26028808

¿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 ☺

CONOCIENDO LA DISTRIBUCIÓN NORMAL

Los valores de la media y la mediana, así como del cuartil 2 si lo calculamos, coinciden o están muy cerca.

Los valores de la media-0.67*desviación estándar coinciden o están muy cercanos a los del cuartil 1; y los valores de la media+0.67*desviación estándar coinciden o están muy cercanos a los del cuartil 3.

Esto se debe a que estamos ante muestras que se comportan como una distribución normal.

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	88.8485786	64.4016751	68.9979312
	Percentil 97.5	93.3002249	64.0649241	69.0029716
	=Media-1.96*Desv. Estándar	37.2293108	18.0451091	60.7282464
	Percentil 2.5	38.3406551	17.2056814	60.0226049
Regla del 2. POSICIÓN del 95% de los datos	=Media+2*Desv. Estándar	89.3753058	64.8747012	69.0823157
	=Media-2*Desv. Estándar	36.7025836	17.5720829	60.6438619

¡Buen trabajo!

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