

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.

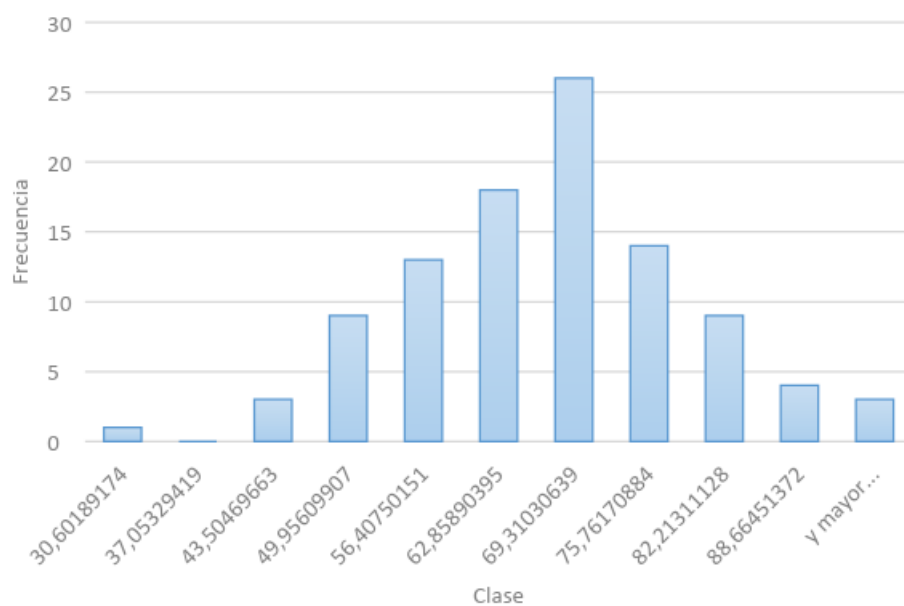
Dibuja los dos histogramas, calcula la media y la desviación estándar de las dos muestras de 100 observaciones y compara las dos distribuciones normales gráficamente y numéricamente.

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

- Muestra m1

CONOCIENDO LA DISTRIBUCIÓN NORMAL

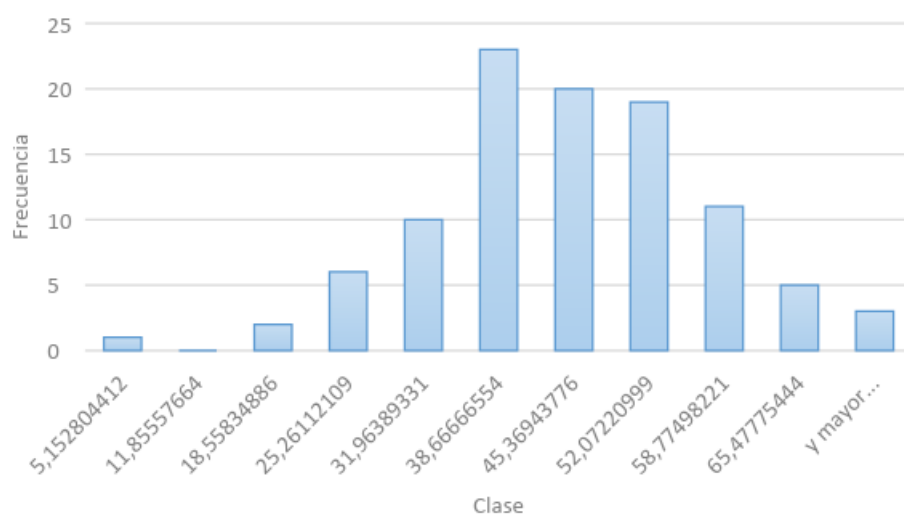
Histograma M1



Media 64,041832
9
Stf 12,434415
7

● Muestra m2

Histograma M2



Media 41,707046
9

CONOCIENDO LA DISTRIBUCIÓN NORMAL

Stf 12,289438
 6

¿Qué observas?

En la M1 que tiene una forma más en pico tenemos un mayor número de observaciones en torno al valor de la media, es más simétrica.

Sin embargo en la m2 que tiene un coeficiente de variación mayor, hay una mayor dispersión de las observaciones.

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.

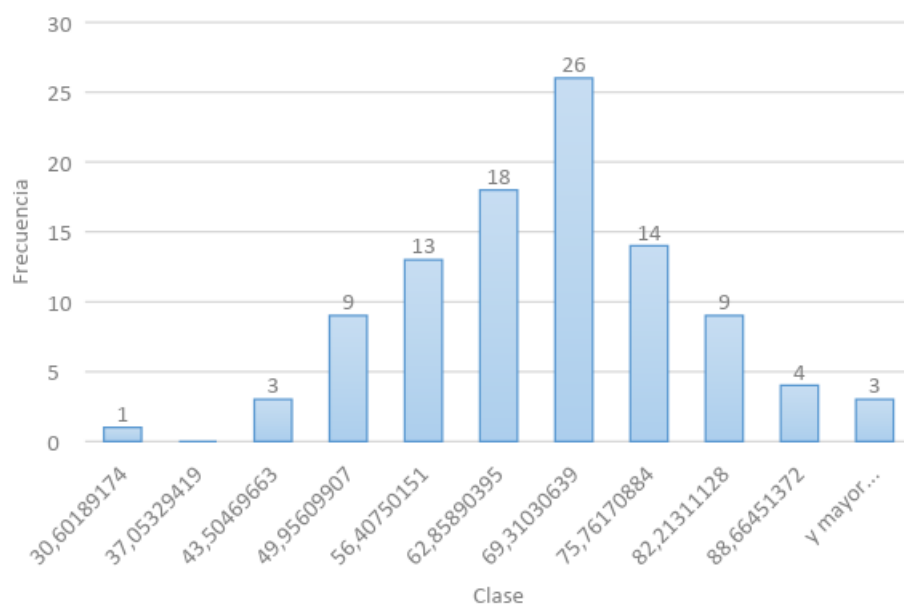
Dibuja los histogramas de la muestra “m1” y “m3”, calcula la media y la desviación estándar de las dos muestras de 100 observaciones y compara las dos distribuciones normales.

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

- Muestra m1

CONOCIENDO LA DISTRIBUCIÓN NORMAL

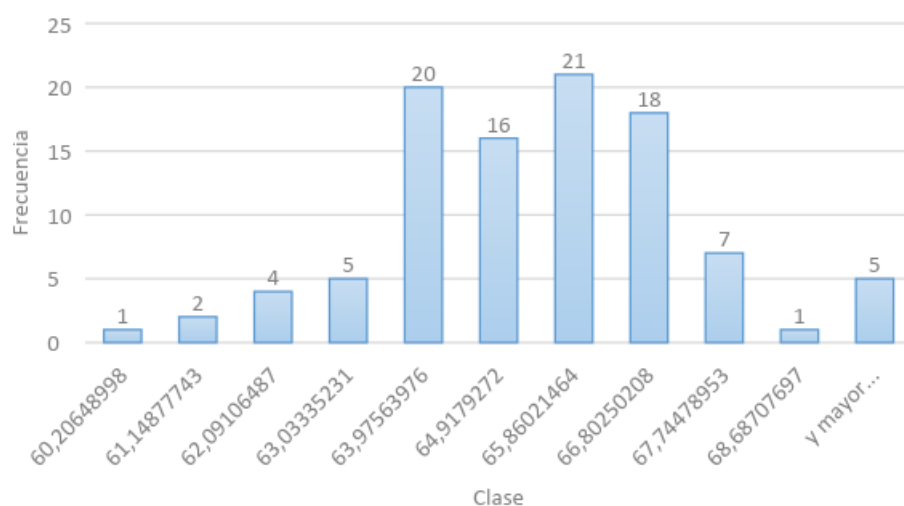
Histograma M1



Media 64,041832
 Stf 12,434415

- Muestra m3

Histograma - M3



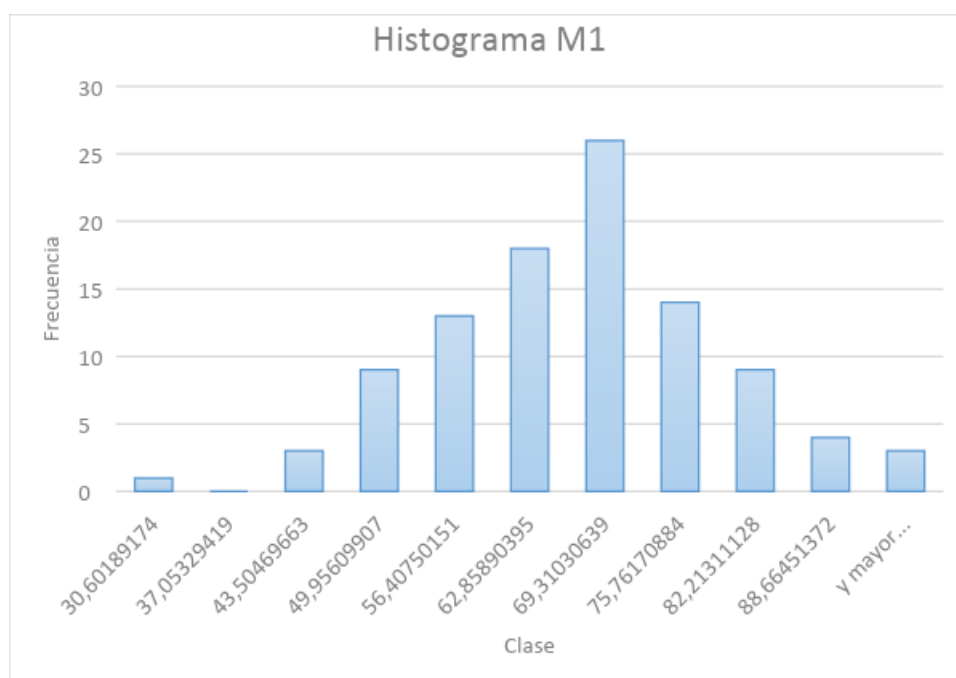
Media 64,985691
 Stf 1,9027041

CONOCIENDO LA DISTRIBUCIÓN NORMAL

¿Qué observas?

Aunque teniendo la $m3$ una desviación mucho menor y sabiendo, por tanto, que la muestra será más homogénea, no se ve claramente representado en los histogramas creados dejando que sea Excel el que determine los rangos de clases. Lo único que sí nos da la pista es el eje de la muestra que se ve claramente que es mucho mayor en la $m1$ (yendo de 30 a más de 88) vs la $m3$ que va de 60 a más de 68.

Para mostrar gráficamente la diferencia de dispersión he representado de nuevo ambas muestras manteniendo el mismo rango de clases de la $m1$.



CONOCIENDO LA DISTRIBUCIÓN NORMAL

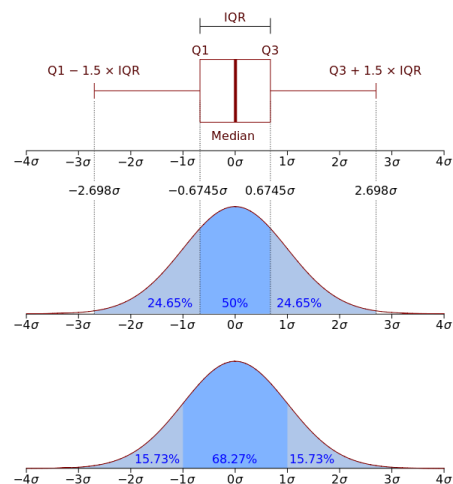


Aplicando las reglas de la distribución normal

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	64,04183292	41,70704689	64,98569097
	Mediana (cuartil 2)	65,79924803	41,77451807	65,06214691
DISPERSIÓN	Desviación Estándar	12,43441574	12,28943859	1,902704175

CONOCIENDO LA DISTRIBUCIÓN NORMAL

POSICIÓN del 50% de los datos más centrados	=Media-0.67*Desv. Estándar	55,71077437	33,47312303	63,71087917
	Cuartil 3	71,06882395	50,24178869	66,41770499
	=Media+0.67*Desv. Estándar	72,37289146	49,94097075	66,26050276
	Cuartil 1	56,29023536	34,70622993	63,5848001
POSICIÓN del 95% de los datos más centrados	=Media+1.96*Desv. Estándar	88,41328777	65,79434653	68,71499115
	Percentil 97.5	91,13548895	70,29548794	69,14391661
	=Media-1.96*Desv. Estándar	39,67037806	17,61974725	61,25639078
	Percentil 2.5	38,56516633	16,68306992	60,875884
FORMA	Curtosis	0,01852667	0,33710407	0,059913268
	Asimetría	-0,012964503	-0,09342867	-0,024826548

¿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 ©

Al tratarse de distribuciones normales, las medidas de posición son conocidas. La media y la mediana son prácticamente iguales en todos los casos.

Y los límites de las variables tanto buscando la concentración del 50% de los datos que prácticamente coinciden con los cuartiles 1 y 3, como la concentración del 95% de los datos que coincide con los percentiles extremos del 97,5 y 2,5.

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"
----------------	--------------	--------------	--------------

CONOCIENDO LA DISTRIBUCIÓN NORMAL

POSICIÓN del 95% de los datos más centrados. Fórmula Exacta	=Media+1.96*Desv. Estándar	88,41328777	65,79434653	68,71499115
	Percentil 97.5	91,13548895	70,29548794	69,14391661
	=Media-1.96*Desv. Estándar	39,67037806	17,61974725	61,25639078
	Percentil 2.5	38,56516633	16,68306992	60,875884
Regla del 2. POSICIÓN del 95% de los datos	=Media+2*Desv. Estándar	88,9106644	66,28592407	68,79109932
	=Media-2*Desv. Estándar	39,17300143	17,12816971	61,18028261

¡Buen trabajo!

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