

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

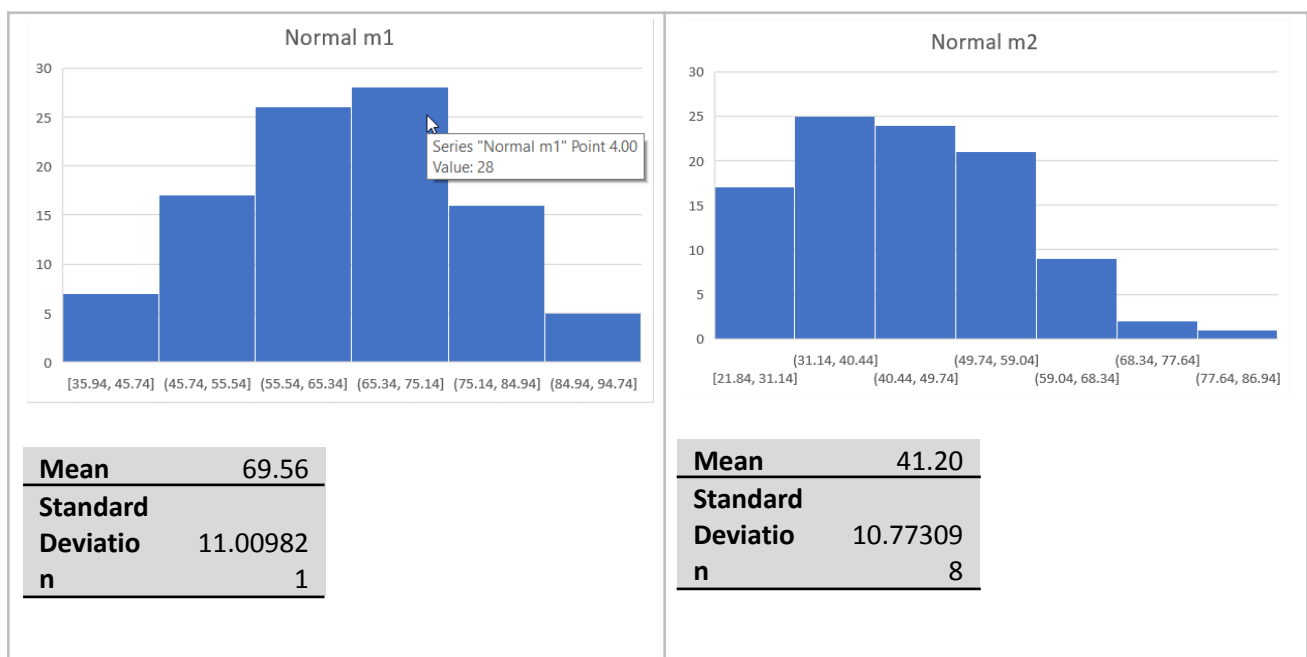
Y otra distribución normal “m2”, con estos dos parámetros:

- 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.



CONOCIENDO LA DISTRIBUCIÓN NORMAL

¿Qué observas?

La desviación estándar es muy similar, pero en m2 los valores se concentran hacia la izquierda por lo que la media es menor que en m1

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.



CONOCIENDO LA DISTRIBUCIÓN NORMAL

¿Qué observas?

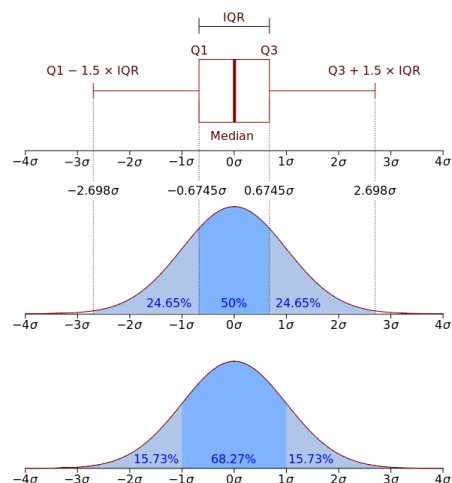
En m1 los valores se distribuyen hacia los lados y la desviación estándar es mayor.
En m3 la desviación estándar es pequeña y los valores se concentran en el centro.

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	69.5553536	41.2033141	64.7394827
	Mediana (cuartil 2)	68.9452783	41.1652864	64.7352294
DISPERSIÓN	Desviación Estándar	11.0098207	10.7730981	1.97545157
POSICIÓN del 50% de los datos más centrados	=Media-0.67*Desv. Estándar	62.1787737	33.9853383	63.4159301
	Cuartil 3	73.3662208	51.6773571	66.2675721
	=Media+0.67*Desv. Estándar	76.9319334	48.4212898	66.0630352
	Cuartil 1	55.5384336	35.3839247	64.1257339
POSICIÓN del 95% de los datos más centrados	=Media+1.96*Desv. Estándar	91.1346021	62.3185864	68.6113678
	Percentil 97.5	90.5082102	67.1151914	59.3590688
	=Media-1.96*Desv. Estándar	47.9761051	20.0880418	60.8675976
	Percentil 2.5	41.8749733	22.6734417	59.3590688
FORMA	Curtosis	-0.1007521	1.39189173	0.18829139
	Asimetría	0.07150722	-0.123756	0.13390581

¿Qué patrón se repite en esta tabla?

Pista: media y mediana deberían ser parecidas, casi iguales

CONOCIENDO LA DISTRIBUCIÓN NORMAL

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

Efectivamente en las tres distribuciones, media y mediana son muy similares. En m3 al ser menor la dispersión, los valores de las posiciones del 50% y el 90% de los datos, presentan menos diferencias.

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	91.134602	62.318586	68.611368
	Percentil 97.5	90.5082102	67.1151914	59.3590688
	=Media-1.96*Desv. Estándar	47.9761051	20.0880418	60.8675976
	Percentil 2.5	41.8749733	22.6734417	59.3590688
Regla del 2. POSICIÓN del 95% de los datos	=Media+2*Desv. Estándar	91.5749949	62.7495103	68.6903858
	=Media-2*Desv. Estándar	47.5357122	19.6571178	60.7885796

¡Buen trabajo!

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