

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:

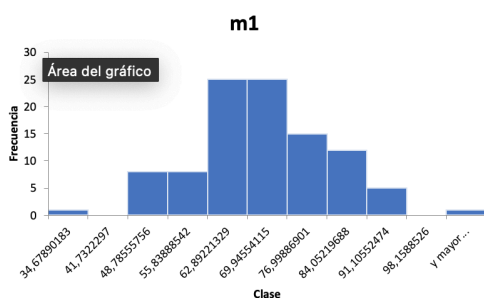
- Valor central (media): 42kg
- Dispersión (desviación estándar): 12 kg

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

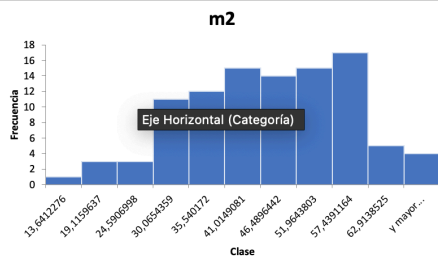


Media = 65,93

Desviación estándar = 11,56

- Muestra m2

CONOCIENDO LA DISTRIBUCIÓN NORMAL



Media = 42,4

Desviación estándar= 12,27

¿Qué observas?

La muestra m1 tiene una centralidad mayor que la muestra m2 y aunque en estos gráficos es difícil de observar, la dispersión es similar aunque la forma sea más aplanada en el caso de m2.

Crea dos distribuciones normales (misma centralidad)

Crea otra distribución normal “m3”, con estos dos parámetros:

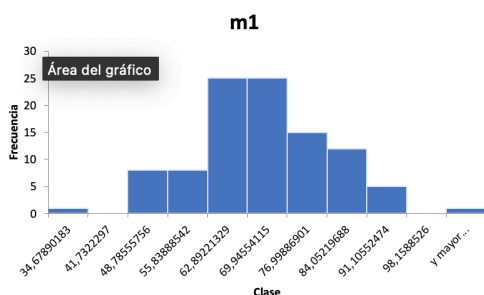
- Valor central (media): 65kg
- Dispersión (desviación estándar): 2 kg

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

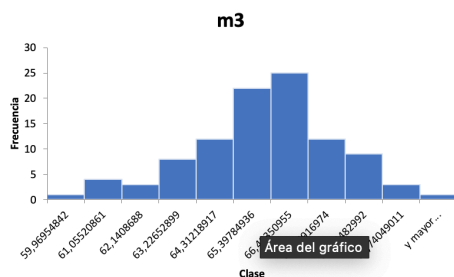


CONOCIENDO LA DISTRIBUCIÓN NORMAL

Media = 65,93

Desviación estándar = 11,56

- Muestra m3



Media = 65,22

Desviación estándar = 2,14

¿Qué observas?

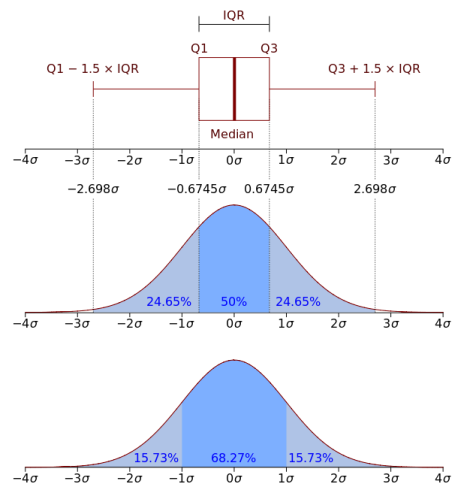
La centralidad es la misma mientras que la dispersión en m3 es mucho menor por lo que los valores se concentran mucho más en torno a la media.

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	65,92542	42,3956594	65,2157693
	Mediana (cuartil 2)	64,8414137	42,3810192	65,3953141
DISPERSIÓN	Desviación Estándar	11,5598196	12,2681936	2,14112836

CONOCIENDO LA DISTRIBUCIÓN NORMAL

POSICIÓN del 50% de los datos más centrados	=Media-0.67*Desv. Estándar	58,1803408	34,1759697	63,7812133
	Cuartil 3	58,3975801	33,2573878	63,7186686
	=Media+0.67*Desv. Estándar	73,6704991	50,6153491	66,6503253
	Cuartil 1	74,025805	49,7194443	66,5370586
POSICIÓN del 95% de los datos más centrados	=Media+1.96*Desv. Estándar	88,5826664	66,4413187	69,4123809
	Percentil 97.5	86,5928112	67,8855216	69,5383385
	=Media-1.96*Desv. Estándar	43,2681735	18,35	61,0191577
	Percentil 2.5	44,5467811	17,3134098	60,3629622
FORMA	Curtosis	0,56615726	-0,274558	0,20039362
	Asimetría	0,21299936	0,16363779	-0,2800762

¿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 crear los valores como una distribución normal, las medidas de centralidad, media y mediana son casi iguales, como en una distribución normal. Además, la posición del 50% de los datos coincide con el rango intercuartílico y la posición del 95% de los datos, con los percentiles 97.5 y 2.5. En cuanto a la forma, pese a que en algunos casos la curtosis o la asimetrían varían de signo, tienden a 0 como en el caso de 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.	=Media+1.96*Desv. Estándar	88,5826664	66,4413187	69,4123809
	Percentil 97.5	86,5928112	67,8855216	69,5383385
	=Media-1.96*Desv. Estándar	43,2681735	18,35	61,0191577

CONOCIENDO LA DISTRIBUCIÓN NORMAL

Fórmula Exacta	Percentil 2.5	44,5467811	17,3134098	60,3629622
Regla del 2. POSICIÓN del 95% de los datos	=Media+2*Desv. Estándar	89,0450592	66,9320465	69,498026
	=Media-2*Desv. Estándar	42,8057807	17,8592723	60,9335126

¡Buen trabajo!

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