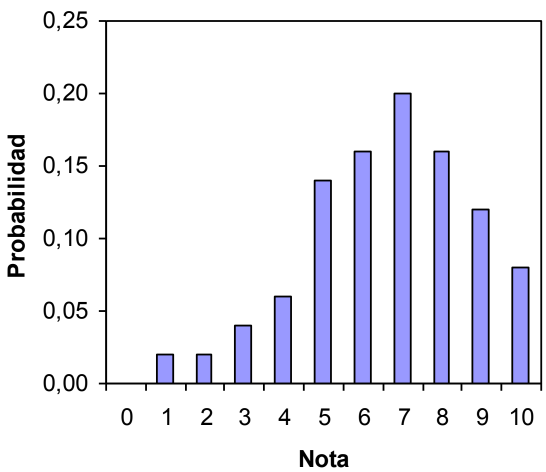
Función de distribución

Cuando se describe la velocidad desde un punto de vista clásico se está pensando en una variable que cambia de forma continua, no discreta, y por tanto, estrictamente, la probabilidad de encontrar un determinado valor exacto de la velocidad es cero. Cuando se tienen variables continuas, su distribución entre la población no se caracteriza por un conjunto discreto de valores de la probabilidad, sino por una función que proporciona la densidad de probabilidad o función de distribución. Piénsese en el caso de un examen de 10 preguntas donde los alumnos pueden obtener, para cada una de ellas, la calificación de 0 ó 1. La nota del examen puede ser por tanto cualquier valor del conjunto discreto [0, 1, 2, ..., 9, 10]. La distribución de la propiedad entre la población de estudiantes (digamos N=50) viene dada por el número de estudiantes que ha obtenido cada uno de los posibles valores [N_0, N_1, ... N_{10}] . Evidentemente, esta distribución puede darse también en forma de probabilidad, sin más que dividir cada uno de los N_i por el número total N.

Nota (x_i)	**N_i**	**p_i**
0	0	0
1	1	0,02
2	1	0,02
3	2	0,04
4	3	0,06
5	7	0,14
6	8	0,16
7	10	0,20
8	8	0,16
9	6	0,12
10	4	0,08
	$\sum_i N_i = 50$	$\sum_i p_i = 1$



La representación de la distribución de la propiedad es en este caso el conjunto de valores de la probabilidad de obtener cada una de las posibles notas en el examen.

En cambio, si se trata de una variable continua no es posible hablar de la probabilidad de encontrar en nuestra población un determinado valor. ¿Cuál es la probabilidad de que una molécula tenga co-

mo módulo de la velocidad 100 m/s?. ¿Qué quiere decir esta pregunta? ¿Exactamente 100,000…00 m/s? Entonces la respuesta sería cero. Cuando se tiene una variable discreta es necesario hablar de la probabilidad de que la variable x tome un determinado valor entre x y x+Δx. Evidentemente, el resultado dependerá de la amplitud del intervalo considerado (Δx).

$$p(x) = \frac{N_{x-x+\Delta x}}{N}$$

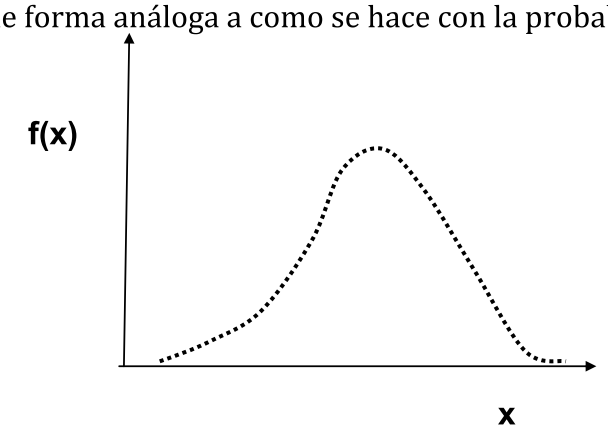
Si nuestra muestra es muy grande se puede obviar este problema considerando intervalos de tamaño infinitesimal y preguntándonos entonces cuál es la probabilidad de que la variable x tome un determinado valor entre x y x+dx. Evidentemente en ese caso la probabilidad será un diferencial, ya que el número de casos presentando valores en ese intervalo sería infinitesimal también (dN):

$$dp(x) = \frac{dN_{x-x+dx}}{N}$$

Para evitar el trabajar con diferenciales y poder tener una función finita se *define* la función de distribución de la variable x (f(x)) como la densidad de probabilidad o probabilidad por unidad de intervalo, de forma que:

$$dp(x) = f(x)dx$$

Teniendo en cuenta que la probabilidad es una magnitud adimensional, la función de distribución tendrá unidades de x^{-1}. Esta función es representable y se puede trabajar con ella de forma análoga a como se hace con la probabilidad.



Así, si estamos interesados en obtener la nota media de nuestro caso anterior se puede simplemente multiplicar cada uno de los posibles resultados (x_i) por su probabilidad y sumar. Esto es equivalente a sumar todas las notas y dividir por el número total de casos (N):

$$\langle x \rangle = \frac{1}{N}\sum_i N_i x_i = \sum_i p_i x_i$$

Pues bien, en el caso de una variable continua, el valor medio también se obtiene multiplicando la variable por la probabilidad (dp(x)) y sumando (integrando) para todos los posibles valores:

$$\langle x \rangle = \int_{\forall x} x\,dp(x) = \int_{\forall x} x f(x)\,dx$$

De igual manera se calcular el valor promedio de cualquier propiedad que dependa de x (h(x)):

$$\langle h(x) \rangle = \int_{\forall x} h(x)\,dp(x) = \int_{\forall x} h(x) f(x)\,dx$$

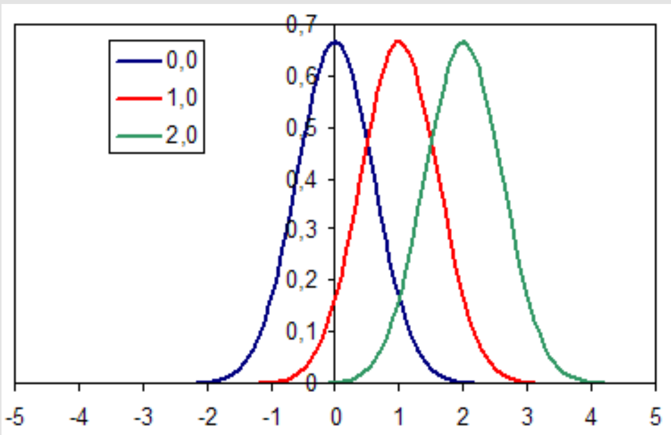
Ejercicio 1.- La distribución gaussiana se utiliza muy frecuentemente en distintos ámbitos para caracterizar funciones de distribución. Su forma genérica es:

$$f(x) = \frac{1}{\left(2\pi\sigma^2\right)^{1/2}} e^{-\frac{(x-\delta)^2}{2\sigma^2}}$$. Representar esta función tomando:

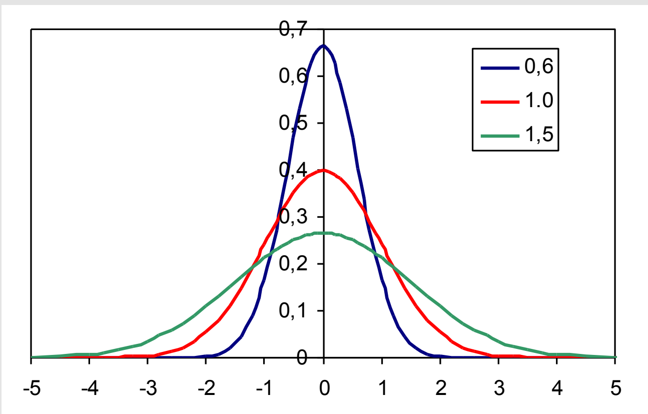
a) σ=0,6 y δ=0, 1 y 2

b) δ=0 y σ=0,6 , 1,0 y 1,5

Solución.- a)



b)



Ejercicio 2.- Al estudiar los ingresos mensuales de los trabajadores de un determinado país se empleó la siguiente función de distribución: $f(x) = Cx^2 e^{-ax^2}$, donde x son los ingresos mensuales en euros y a se determinó que valía 3,785x10^{-6} euros^{-2}.

a) Calcular C sabiendo que la función de distribución debe estar normalizada.

b) ¿Cuáles son los ingresos mensuales medios de un habitante de ese país?

c) Representar la función de distribución. Indicar como determinaría la proporción de habitantes del país que tienen ingresos mensuales menores que el valor medio? ¿y mayores?.

Solución.-

a) $$\int_{\forall x} dp_x = \int_{\forall x} f(x)dx = 1 = \int_0^\infty Cx^2 e^{-ax^2} dx = C\frac{1}{4a}\frac{\pi^{1/2}}{a^{1/2}}$$

$$1 = \int_0^\infty Cx^2 e^{-ax^2} dx = C\frac{1}{4a}\frac{\pi^{1/2}}{a^{1/2}} \text{ (recordando que } \int_0^\infty x^{2n} e^{-ax^2} dx = \frac{(2n)!\pi^{\frac{1}{2}}}{2^{2n+1} n! a^{n+\frac{1}{2}}} \text{)}$$

de donde $C = 4\dfrac{a^{3/2}}{\pi^{1/2}} = 1{,}6618\text{x}10^{-8}$ euros^{-3} .

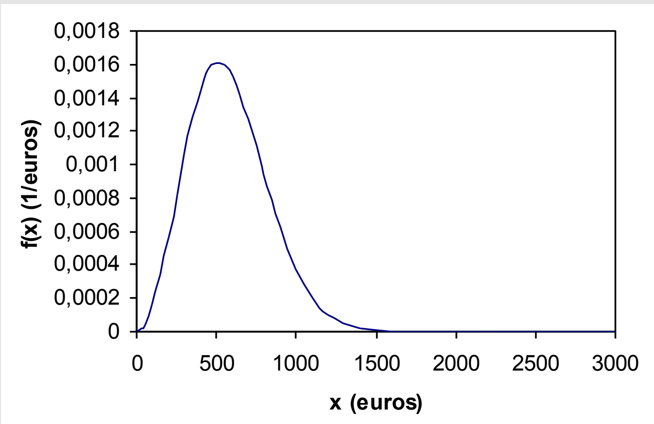
b) Recordando que $\langle x \rangle = \int_{\forall x} x dp_x = \int_{\forall x} x f(x)dx$,

$$\langle x \rangle = \int_0^\infty xCx^2 e^{-ax^2} dx = C\int_0^\infty x^3 e^{-ax^2} dx = \frac{4a^{3/2}}{\pi^{1/2}}\int_0^\infty x^3 e^{-ax^2} dx .$$

Como $\int_0^\infty x^{2n+1} e^{-ax^2} dx = \dfrac{n!}{2a^{n+1}}$, de la solución de la integral superior se obtiene:

$$\langle x \rangle = \frac{4a^{3/2}}{\pi^{1/2}}\frac{1}{2a^2} = \frac{2}{(\pi a)^{1/2}} = 580 \quad \text{euros}$$

c)



$$\frac{N(x<\langle x\rangle)}{N}=\int_0^{\langle x\rangle}\frac{dN_x}{N}=\int_0^{\langle x\rangle}dp_x=\int_0^{\langle x\rangle}f(x)dx \quad ;$$

$$\frac{N(x>\langle x\rangle)}{N}=\int_{\langle x\rangle}^{\infty}\frac{dN_x}{N}=\int_{\langle x\rangle}^{\infty}dp_x=\int_{\langle x\rangle}^{\infty}f(x)dx$$

$$\text{con} \quad \int_0^{\langle x\rangle}dp_x+\int_{\langle x\rangle}^{\infty}dp_x=1$$