

II.4 Distribuciones Continuas Habituales

March 31, 2019

Variable Aleatoria Gaussiana y otras habituales

Variable Aleatoria Uniforme

Denominamos **soporte de la variable aleatoria** a los valores de la misma en los que la función de densidad de probabilidad toma valores finitos distintos de cero. Considerando un intervalo de soporte $I = [a, b]$, se llama **variable aleatoria uniforme** a aquella en que todos los intervalos de igual longitud y totalmente contenidos en el intervalo de soporte sean **equiprobables**. La variable aleatoria uniforme tiene una función de densidad de probabilidad:

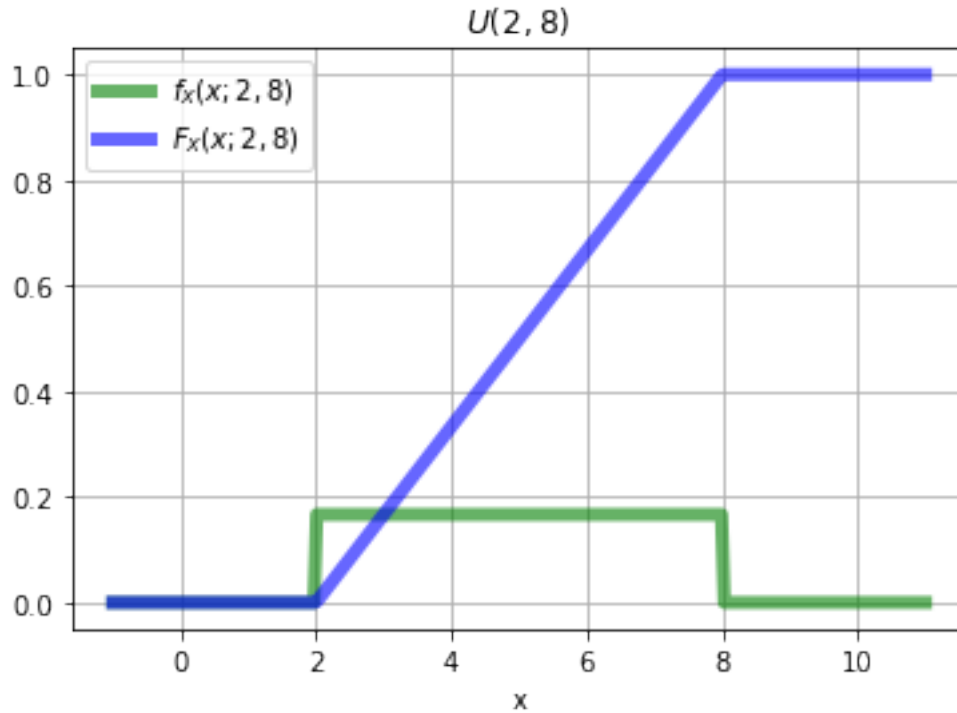
$$U(a, b) \equiv f_X(x; a, b) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{resto} \end{cases}$$

Y función de distribución de probabilidad:

$$F_X(x; a, b) = \begin{cases} 0 & \text{si } x < a \\ \frac{1}{b-a}(x-a) & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

Calculando las integrales se obtienen fácilmente la media y la varianza:

- $\eta_X = E(X) = \int_{-\infty}^{\infty} x f_X(x; a, b) = \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2}$
- $E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x; a, b) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$
- $\sigma_X^2 = Var(X) = E((X - \eta_X)^2) = E(X^2) - \eta_X^2 = \frac{(b-a)^2}{12}$



Funciones Condicionadas

Consideremos el caso ya visto en que el suceso condicionante A podemos definirlo con intervalos de la variable aleatoria. Suponiendo que $A = \{a \leq x_1 \leq X \leq x_2 \leq b\}$. Adviértase que si $x_1 < a$ ó si $x_2 > b$, al extenderse el suceso condicionante fuera del soporte de la distribución, A podría escribirse de forma equivalente con $x_1 = a$ ó $x_2 = b$, respectivamente.

$$F_X(x / \{x_1 \leq X \leq x_2\}) = \frac{P(\{X \leq x \cap \{x_1 \leq X \leq x_2\}\})}{P(\{x_1 \leq X \leq x_2\})} =$$

$$= \begin{cases} 0 & x < x_1 \\ \frac{F_X(x) - F_X(x_1^-)}{F_X(x_2) - F_X(x_1^-)} & x_1 \leq x \leq x_2 \\ 1 & x > x_2 \end{cases}$$

$$f_X(x / \{x_1 \leq X \leq x_2\}) = \begin{cases} 0 & x < x_1 \\ \frac{f_X(x)}{F_X(x_2) - F_X(x_1^-)} & x_1 \leq x \leq x_2 \\ 0 & x > x_2 \end{cases}$$

Sustituyendo para el caso concreto de la distribución uniforme, es fácil obtener:

$$f_X(x/A) = \begin{cases} \frac{1}{x_2 - x_1} & \text{si } x_1 \leq x \leq x_2 \\ 0 & \text{resto} \end{cases}$$

$$F_X(x/A) = \begin{cases} 0 & \text{si } x < x_1 \\ \frac{1}{x_2 - x_1}(x - x_1) & \text{si } x_1 \leq x \leq x_2 \\ 1 & \text{si } x > x_2 \end{cases}$$

Ejemplo

Considérese que un suceso acontece con certeza en un intervalo, pero de forma totalmente aleatoria dentro del mismo. Piénsese, por ejemplo, en la llamada de una persona concreta que siempre se recibe diariamente entre 22h y 23h.

Se pide:

1. Modelar la VA X que representa el momento de recibir la llamada, obteniendo tanto la fdp como la FDP
2. Calcular la media y la varianza
3. Calcular la probabilidad de recibir la llamada entre las 22h y las 22h:15m
4. Calcular la probabilidad de recibir la llamada justo a las 22h:15m
5. Si hasta las 22h:15m no se ha recibido la llamada, ¿cuál será la probabilidad de recibirla entre las 22h:45m y las 23h?

Variable Aleatoria Exponencial

La variable aleatoria exponencial modela el tiempo o espacio entre el acontecimiento de dos sucesos sucesivos en un Proceso Puntual de Poisson. Tiene la siguiente función de densidad de probabilidad:

$$\exp(\lambda) \equiv f_X(x; \lambda) = \lambda e^{-\lambda x} u(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases}$$

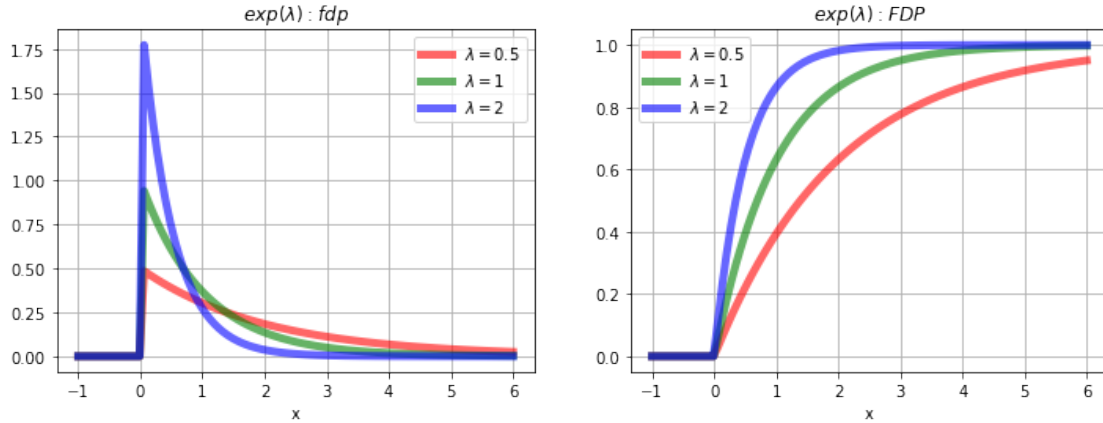
Donde $u(x)$ es la **función escalón**. La función de distribución de probabilidad es:

$$F_X(x; \lambda) = (1 - e^{-\lambda x}) u(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

$\lambda > 0$ se denomina **tasa**, si bien a veces se utiliza la **escala** $\beta = \lambda^{-1}$.

Calculando las integrales por partes se obtienen la media y la varianza:

- $\eta_X = E(X) = \int_{-\infty}^{\infty} x f_X(x; \lambda) = \lambda \int_0^{\infty} x e^{-\lambda x} dx = \frac{1}{\lambda} = \beta$
- $E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x; \lambda) = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = \frac{2}{\lambda^2} = 2\beta^2$
- $\sigma_X^2 = Var(X) = E((X - \eta_X)^2) = E(X^2) - \eta_X^2 = \frac{1}{\lambda^2} = \beta^2$



Funciones condicionadas y propiedad sin memoria

Considérese el suceso condicionante $A = \{X > x_1\}$. Las probabilidades condicionadas se obtienen haciendo $x_2 = \infty$ en las ecuaciones ya vistas:

$$F_X(x/X > x_1) = \frac{F_X(x) - F_X(x_1)}{1 - F_X(x_1)} u(x - x_1)$$

$$f_X(x/X > x_1) = \frac{f_X(x)}{1 - F_X(x_1)}$$

Sustituyendo las expresiones correspondientes a la distribución exponencial, resulta:

$$F_X(x/X > x_1) = \left(1 - e^{-\lambda(x-x_1)}\right) u(x - x_1) = F_X(x - x_1)$$

$$f_X(x/X > x_1) = e^{-\lambda(x-x_1)} u(x - x_1) = f_X(x - x_1)$$

Como puede verse, **la distribución exponencial no tiene memoria**, esto es, transcurrido un lapso x_1 desde el último suceso la distribución se comporta como si acabara de ocurrir, desplazándose hasta x_1 .

Ejemplo

Considérese que un servidor informático recibe peticiones independientes de clientes con una tasa de llegada $\lambda = 2$ cada segundo. La llegada de peticiones a lo largo del tiempo puede modelarse como un Proceso Puntual de Poisson por ser las llegadas independientes.

- Si acaba de recibirse una petición, ¿cuál será la probabilidad de recibir otra en el próximo medio segundo?
- ¿Y cuál será la probabilidad de no recibir otra hasta dentro de un segundo?
- Tras la última llegada de una petición, extrañamente no ha entrado ninguna durante diez segundos. ¿Cuál será la probabilidad de que lo haga en el medio segundo siguiente?

La Variable Aleatoria Gaussiana

Distribución Normal Estándar

Consideremos la función $\varphi(x)$, definida como sigue:

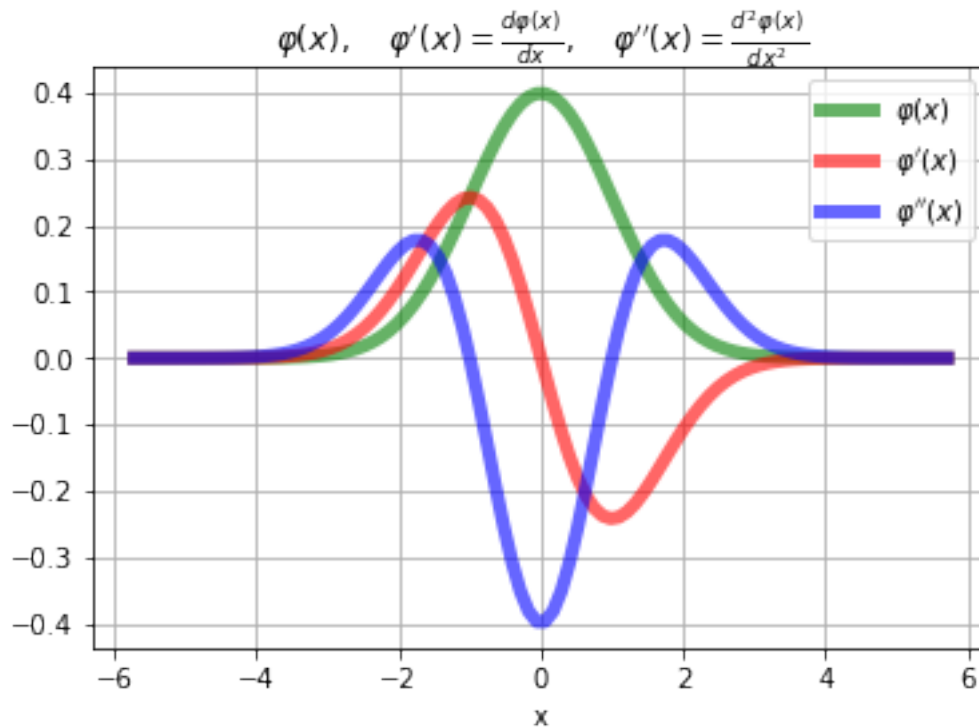
$$\varphi_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Puede advertirse que se trata de una función de densidad de probabilidad, pues:

$$\varphi(x) > 0 \quad \forall x \in \mathbb{R} \quad \int_{-\infty}^{\infty} \varphi(x) dx = 1$$

Analicemos la función $\varphi_X(x)$, que denominamos **Normal Estándar**:

- Es una función par pues $\varphi_X(x) = \varphi_X(-x)$
- $\varphi'_X(x) = -x\varphi_X(x) \implies \varphi'_X(0) = 0$
- $\varphi''_X(x) = (x^2 - 1)\varphi_X(x) \implies \varphi''_X(-1) = \varphi''_X(1) = 0$
- Máximo: $\varphi_X(0) = \frac{1}{\sqrt{2\pi}} \approx 0.399$. Puntos de inflexión: $\varphi_X(-1) = \varphi_X(1) = \frac{1}{\sqrt{2\pi e}} \approx 0.242$



La **función de distribución normal estándar** es:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$$

En *Teoría de la Comunicación* suele utilizarse la función $Q(x)$, que es el complemento a 1 de $\Phi(x)$:

$$Q(x) = 1 - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du$$

En ingeniería y ciencias es habitual utilizar la **función error**, $erf(x)$:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-u^2} du = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{2}x}^{\sqrt{2}x} e^{-\frac{u^2}{2}} du$$

que satisface:

- Función impar: $erf(x) = -erf(-x)$
- $erf(\infty) = 1$
- $erf(1) \approx 0.8427$

Su complemento a 1 es la **función error complementaria**:

$$erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-u^2} du = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2}x}^\infty e^{-\frac{u^2}{2}} du$$

Pueden encontrarse fácilmente relaciones entre las funciones anteriores:

$$\Phi(x) = \frac{1}{2} \left[1 + \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du \right] = \frac{1}{2} \left[1 + erf\left(\frac{x}{\sqrt{2}}\right) \right]$$

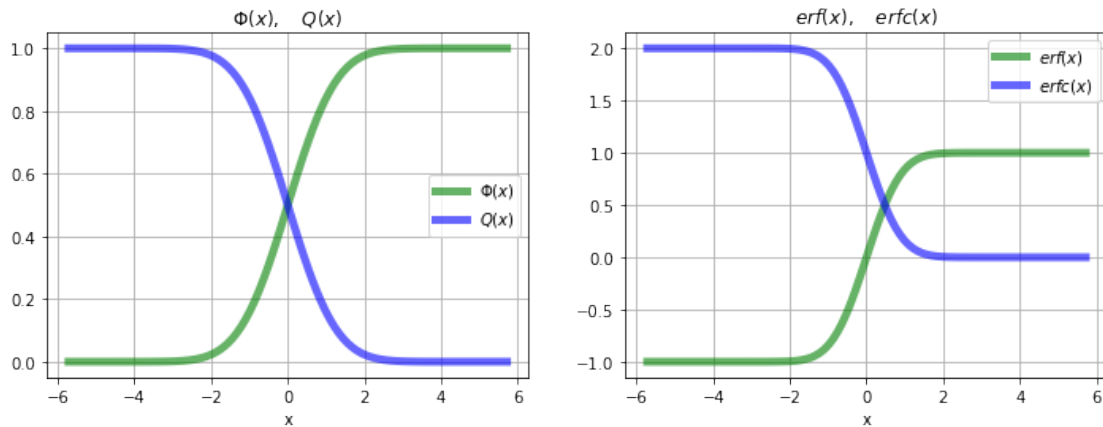
$$Q(x) = 1 - \Phi(x) = \frac{1}{2} \left[1 - erf\left(\frac{x}{\sqrt{2}}\right) \right] = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)$$

Y puede calcularse la probabilidad de los intervalos:

$$P(x_1 < X \leq x_2) = \Phi(x_2) - \Phi(x_1) = \frac{1}{2} erf\left(\frac{x_2}{\sqrt{2}}\right) - \frac{1}{2} erf\left(\frac{x_1}{\sqrt{2}}\right)$$

En el caso que el intervalo esté centrado en el origen $x_1 = -x_2$:

$$P(-x_2 < X \leq x_2) = \Phi(x_2) - \Phi(x_1) = erf\left(\frac{x_2}{\sqrt{2}}\right)$$



Puede calcularse fácilmente la concentración de masa de probabilidad en torno al origen:

- $P(|X| \leq 1) = \text{erf}(\frac{1}{\sqrt{2}}) \approx 0.682689492137086$
- $P(|X| \leq 2) = \text{erf}(\frac{2}{\sqrt{2}}) \approx 0.954499736103642$
- $P(|X| \leq 3) = \text{erf}(\frac{3}{\sqrt{2}}) \approx 0.997300203936740$

Adviértase, que al ser la fdp normal estándar una función par, $\varphi(x) = \varphi(-x)$, la esperanza matemática resulta nula. El cálculo de la varianza requiere calcular la integral, y es igual a 1:

- $\eta_X = E(X) = \int_{-\infty}^{\infty} x\varphi(x)dx = 0$
- $E(X^2) = \int_{-\infty}^{\infty} x^2\varphi(x)dx = 1$
- $\sigma_X^2 = \text{Var}(X) = E(X^2) - E^2(X) = 1 \implies \sigma_X = 1$

Por tanto, las concentraciones de masa de probabilidad anteriores están calculadas a una, dos y tres σ de la media, respectivamente.

Distribución Gaussiana o Normal

Si la variable aleatoria X tiene una distribución Normal Estándar $\varphi(x)$ (media nula y varianza unidad), podemos generar una nueva variable aleatoria con media η_Y y varianza σ_Y^2 sin más que aplicar la transformación lineal:

$$Y = \sigma_Y X + \eta_Y$$

Puede demostrarse que la variable aleatoria Y tiene la siguiente distribución:

$$F_Y(y) = \Phi\left(\frac{y - \eta_Y}{\sigma_Y}\right)$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{\sigma_Y} \varphi\left(\frac{y}{\sigma_Y}\right)$$

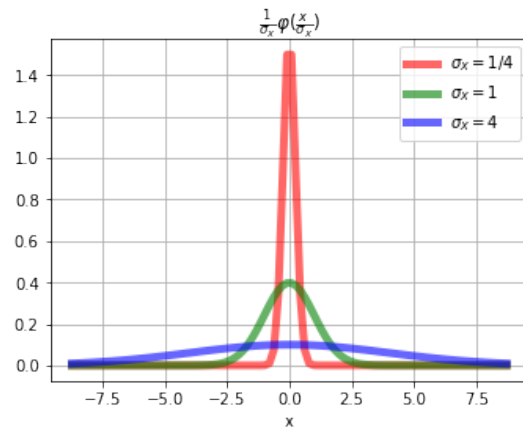
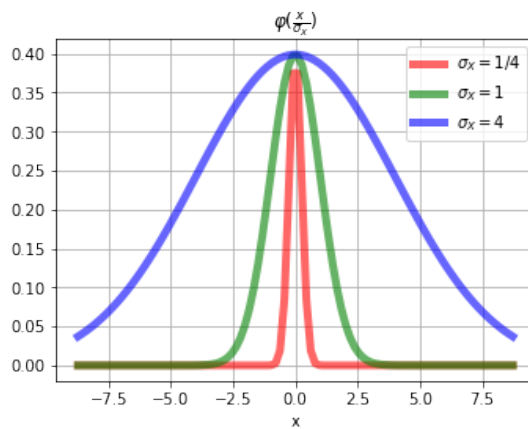
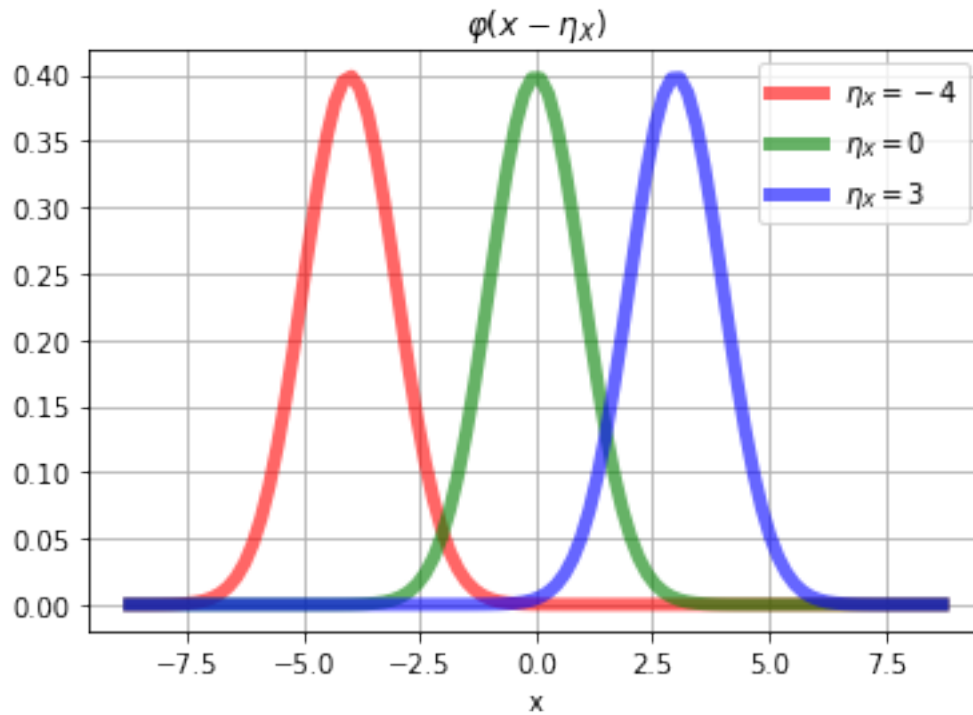
Intentemos ganar cierta intuición trabajando con estas transformaciones de la función de densidad Normal Estándar. Adviértase que, en lo sucesivo, tras los cambios de variable volvemos a llamar a la variable X . Aunque al principio puede resultar confuso, es habitual trabajar así para tener las funciones de interés con x de variable independiente.

La función $\varphi(x)$ puede desplazarse en el eje real, y seguirá siendo una función de densidad de probabilidad:

$$\varphi(x - \eta_X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x - \eta_X)^2}$$

Sin embargo, si le aplicamos un cambio de escala $\varphi(\frac{x}{\sigma_X})$ puede advertirse que el área bajo la curva es σ_X . Por ello, es necesario normalizar por el área resultante tras el cambio de escala, para seguir teniendo una función de densidad de probabilidad:

$$\frac{1}{\sigma_X} \varphi\left(\frac{x}{\sigma_X}\right)$$



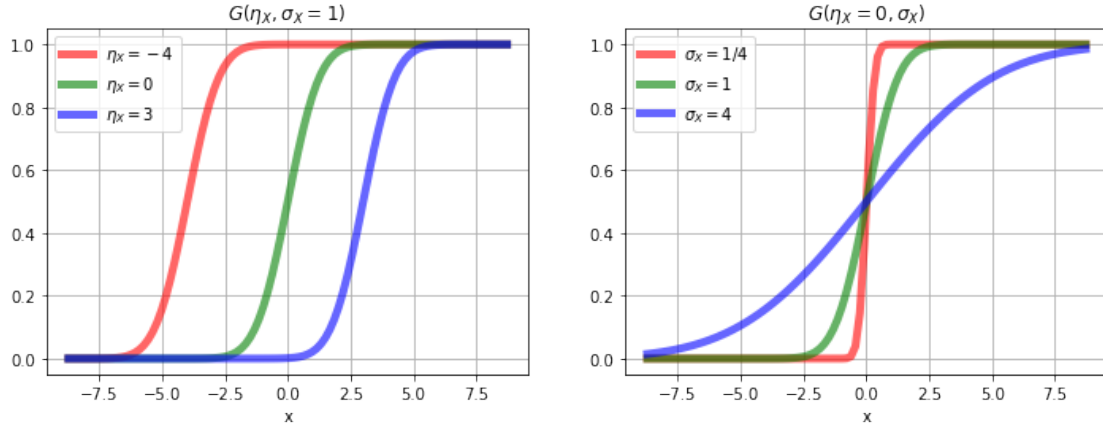
Se define la distribución normal de parámetros η_X y σ_X a la distribución de probabilidad con la siguiente función de densidad:

$$N(\eta_X, \sigma_X) \equiv f_X(x; \eta_X, \sigma_X) = \frac{1}{\sigma_X} \varphi\left(\frac{x - \eta_X}{\sigma_X}\right) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x - \eta_X)^2}{2\sigma_X^2}}$$

Como se ha visto, $E(X) = \eta_X$ y $\text{Var}(X) = \sigma_X^2$.
La función de distribución de probabilidad es:

$$G(\eta_X, \sigma_X) \equiv F_X(x; \eta_X, \sigma_X) = \Phi\left(\frac{x - \eta_X}{\sigma_X}\right) = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^x e^{-\frac{(u - \eta_X)^2}{2\sigma_X^2}} du =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{\frac{x - \eta_X}{\sigma_X}} e^{-\frac{u^2}{2}} du = \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{x - \eta_X}{\sigma_X \sqrt{2}}\right) \right]$$



Propiedades asintóticas

Aproximación de la Distribución Binomial Conforme al **Teorema de Moivre-Laplace** la distribución Binomial $B(n, p)$ puede aproximarse por la distribución Normal $N(np, \sqrt{np(1-p)})$ cuando n es grande y p no está cerca de 0 ni de 1. El Teorema prueba la convergencia en el límite. En la práctica, se adopta la aproximación cuando $n > 20$ y $p \approx 0.5$. Cuanto mayor sea n más puede apartarse p del valor 0.5, aunque nunca debe aproximarse a 0 ni 1.

$$B(n, p) \approx N(np, \sqrt{np(1-p)})$$

Aproximación de la Distribución de Poisson La distribución de Poisson $Pois(\lambda)$ puede aproximarse por la distribución Normal $N(\lambda, \lambda)$ cuando λ es grande

$$Pois(\lambda) \approx N(\lambda, \lambda)$$

Función de densidad condicionada

Consideremos ahora el suceso condicionante $A = \{x_1 < X \leq x_2\}$.

$$F_X(x / \{x_1 < X \leq x_2\}) = \begin{cases} 0 & x \leq x_1 \\ \frac{F_X(x) - F_X(x_1)}{F_X(x_2) - F_X(x_1)} & x_1 < x \leq x_2 \\ 1 & x > x_2 \end{cases}$$

$$f_X(x / \{x_1 < X \leq x_2\}) = \begin{cases} 0 & x \leq x_1 \\ \frac{f_X(x)}{F_X(x_2) - F_X(x_1)} & x_1 < x \leq x_2 \\ 0 & x > x_2 \end{cases}$$

Sustituyendo para la variable aleatoria Gaussiana:

$$F_X(x_2) - F_X(x_1) = \frac{1}{2} \operatorname{erf}\left(\frac{x_2 - \eta_X}{\sigma_X \sqrt{2}}\right) - \frac{1}{2} \operatorname{erf}\left(\frac{x_1 - \eta_X}{\sigma_X \sqrt{2}}\right)$$

Si x_1 y x_2 son simétricos respecto de η_X :

$$F_X(x_2) - F_X(x_1) = \operatorname{erf}\left(\frac{x_2 - \eta_X}{\sigma_X \sqrt{2}}\right)$$

Adviértase que, considerando este factor normalizante, podemos reutilizar las expresiones de la distribución total (no condicionada) para el cálculo de probabilidades en los subintervalos incluidos en el suceso condicionante A .

Mezcla de Gaussianas

Es habitual en la práctica confrontar distribuciones multimodales. Con frecuencia se considera que estamos ante una superposición de distribuciones Gaussianas, una por cada una de las modas existentes, con sus respectivos valores medios y varianzas.

A veces conviene considerar que cada una de tales distribuciones representa una subpoblación, y que cada observación de la variable aleatoria proviene de una de ellas. Asociemos a cada subpoblación con una hipótesis H_i de la que conocemos su probabilidad *a priori*, $P(H_i)$. Consideremos que también conocemos las verosimilitudes de las observaciones en relación a las subpoblaciones, que podemos modelar mediante las probabilidades condicionadas:

$$f_{X_i}(x/H_i) \equiv N(\eta_{X_i}, \sigma_{X_i})$$

Podemos utilizar el Teorema de la Probabilidad Total para obtener la fdp de la **mezcla de Gaussianas**:

$$f_X(x) = f_{X_1}(x/H_1)P(H_1) + \dots + f_{X_I}(x/H_I)P(H_I)$$

Ejemplo

En una población, el 48% de las personas son hombres y el resto mujeres. Las estaturas de cada subpoblación se modela con una distribución Gaussiana, de la siguiente manera. La estatura media de las mujeres es de 165 cm y su desviación típica de 8 cm, mientras que la estatura media de los hombres es de 175 cm y su desviación típica es de 10 cm.

- Calcule cuál es la probabilidad de que un hombre mida más de 2 m. Igualmente con una mujer
- Considere ahora subpoblaciones de baloncestistas. En la femenina todas las estaturas son superiores a 170 cm, mientras que en las masculinas son todas superiores a 180 cm. ¿Cuál es la probabilidad de que una jugadora mida más de 2 m? ¿Y un jugador?

- Suponga que tenemos una persona cuyo género desconocemos y que mide 170 cm. ¿Qué es más verosímil, que sea un hombre o una mujer?
- Obtenga utilizando el Teorema de la Probabilidad Total la mezcla de Gaussianas que modela la distribución total de hombres y mujeres
- Utilice el Teorema de Bayes para ver qué es más probable, que la persona de 170 cm sea un hombre o una mujer.
- ¿Cuál es la probabilidad de error si decide que es un hombre? ¿Y si decide que es una mujer?
- Elabore una regla de decisión que permita seleccionar el género por estatura minimizando el error. ¿Cuál será la probabilidad de error de esta regla?