1 Distribución normal

La distribución normal (o gaussiana) juega un papel central en la teoría de la probabilidad y en la estadística. Empíricamente, muchas cantidades observables en el mundo real, como la altura, el peso, los grados y el error de medición, a menudo están bien aproximadas por la distribución normal. Además, el teorema del límite central establece que, en condiciones muy generales, la distribución del promedio de un gran número de variables aleatorias independientes converge a la distribución normal.

En esta notas mostraremos las propiedades básicas de las variables aleatorias normales univariadas.

1.1 Distribución normal estándar

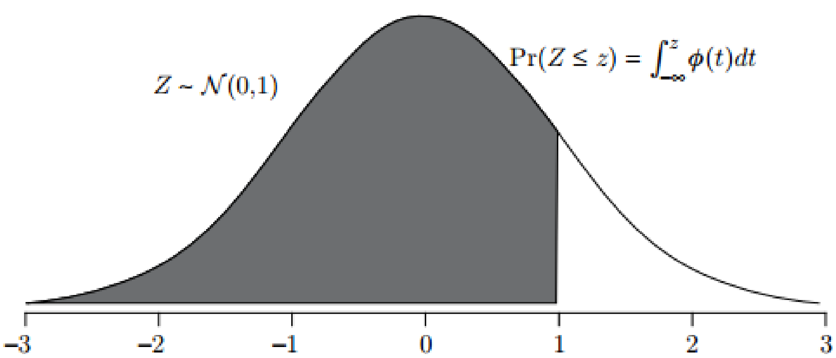
La distribución normal estándar, denotada como $N(0,1)$, es el componente básico de todas las distribuciones normales univariadas y multivariadas.

Decimos que una variable aleatoria Z es normal estándar si es $N(0,1)$. La distribución normal estándar es una distribución continua sobre los números reales, con función de densidad:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2},$$

y una función de distribución:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx.$$



Si bien la función de distribución $\Phi(z)$ está bien definida para cualquier z, no tiene una expresión de forma cerrada. Los valores reales de la distribución normal estándar se pueden calcular numéricamente.

La distribución está relacionada con lo que se conoce como la función de error en estadística, que comúnmente se denota por:

$$\text{erf}(z) = \frac{2}{\sqrt{2\pi}} \int_{0}^{z} e^{-x^2/2} dx.$$

Es simple verificar:

$$\Phi(z) = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{z}{\sqrt{2}}\right).$$

Como se mencionó los valores reales de la distribución normal estándar se pueden calcular numéricamente, como se ve en la siguiente tabla:

Table of Distribution Function Φ
of the standard normal $\mathbf{N}(0,1)$ distribution

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993

Para $z < 0$ se usa $\Phi(z) = 1 - \Phi(-z)$.

Desde que la densidad de la distribución normal estándar Z es simétrica con respecto a $z = 0$, se sigue que $\mathbb{E}(Z) = 0$. La varianza de Z puede ser calculado usando integración por partes (donde las partes usadas abajos son $u = x$ y $dv = xe^{x^2/2}dx$:

$$\begin{aligned}
\mathrm{Var}(Z) = \mathbb{E}(Z^2) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Big(x\Big)\Big(xe^{-x^2/2}\Big) dx \\
&= -\frac{1}{\sqrt{2\pi}} xe^{-x^2/2}|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1.
\end{aligned}$$

En la última igualdad, el primer término es 0, y ya hemos observado que el segundo término es igual a 1.

1.2 Distribución normal univariada general

La distribución normal univariada se caracteriza por dos parámetros μ y σ, que corresponden a la media y a la desviación estándar, y se denota por $N(\mu, \sigma^2)$. Ten en cuenta que usamos la varianza en lugar de la desviación estándar en la forma en que denotamos la distribución normal.

Como antes, podemos decir que X es una variable aleatoria normal (con los parámetros μ y σ) para denotar que X es una variable aleatoria con una distribución normal.

La función de de densidad de una variable aleatoria normal $X \sim N(\mu, \sigma^2)$ es:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-((x-\mu)/\sigma)^2/2}, \tag{1}$$

y su función de disribución es:

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-((t-\mu)/\sigma)^2/2} dt.$$

Estas expresiones generalizan las expresiones correspondientes para las variables aleatorias normales estándar, donde $\mu = 0$ y $\sigma = 1$.

En efecto, sea X una variable aleatoria normal con parámetros μ y σ y sea $Z = (X - \mu)/\sigma$. Entonces:

$$\mathbb{P}(Z \leq z) = \mathbb{P}(X \leq \sigma z + \mu) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\sigma z + \mu} e^{-((t-\mu)/\sigma)^2/2} dt$$

Sustituyendo $x = (t - \mu)/\sigma$ y usando $dt = \sigma dx$, podemos encontrar la densidad de la distribución normal,

$$\mathbb{P}(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx = \Phi(z).$$

Vemos que la distribución normal X es una transformación lineal de la distribución normal estándar. Es decir, si X es una variable aleatoria normal con los parámetros μ y σ, entonces $Z = (X - \mu)/\sigma$ es una variable aleatoria normal estándar (con media 0 y varianza 1) y de manera similar, si Z es un estándar normal variable aleatoria, entonces $X = \sigma Z + \mu$ es una variable aleatoria normal con los parámetros μ y σ.

Hemos mostrado lo siguiente:

Lema 1.1 Una variable aleatoria tiene una distribución normal si y solo si es una transformación lineal de una variable aleatoria normal estándar.

Dado que una variable aleatoria X con una distribución $N(\mu, \sigma^2)$ tiene la misma distribución que $\sigma z + \mu$, tenemos que $\mathbb{E}(X) = \mu$ y $\text{Var}(X) = \sigma^2$, donde μ y σ son la media y desviación estándar.

1.3 Ejemplo: Detección de señales

Supongamos que tenemos un transmisor que está enviando un bit a través de la codificación $S \in \{-1, +1\}$ y el canal de comunicación agrega ruido Y a la señal, donde Y es una variable aleatoria normal con media 0 y desviación estándar σ. El receptor decodifica tomando el signo de la señal recibida, $R = \text{sgn}(S + Y)$. (Aquí sgn es la función signo, donde $\text{sgn}(x) = 1$ si $x > 0$, $\text{sgn}(x) = -1$ si $x < 0$ y por convención $\text{sgn}(0) = 0$.)

Encontremos la probabilidad de que el mensaje decodificado es diferente del mensaje original. Esa es la probabilidad de que $R \neq S$.

La probabilidad de un error cuando $S = 1$ es la probabilidad que $Y \leq -1$:

$$\mathbb{P}(Y \leq -1) = \mathbb{P}\left(\frac{Y - \mu}{\sigma} \leq \frac{-1 - \mu}{\sigma}\right) = \Phi\left(-\frac{1}{\sigma}\right).$$

De manera similar, la probabilidad de un error cuando $S = -1$ es la probabilidad que $Y \geq -1$:

$$\mathbb{P}(Y \geq 1) = 1 - \mathbb{P}\left(\frac{Y - \mu}{\sigma} \leq \frac{1 - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{1}{\sigma}\right).$$

Por la simetría de la distribución normal alrededor de la media, $\Phi\left(-\frac{1}{\sigma}\right) = 1 - \Phi\left(\frac{1}{\sigma}\right)$ y el error total es: $2\left(1 - \Phi\left(\frac{1}{\sigma}\right)\right)$.

Podemos usar la tabla anterior para determinar los errores de probabilidad. Para $\sigma = 0.5, 1, 2$, encontramos $\Phi(2) = 0.9772, \Phi(1) = 0.8413$ y $\Phi(0.50) = 0.6915$.

Los errores de probabilidades son por tanto $0.0456, 0.3174$ y 0.6170 respectivamente.

1.4 Funciones generadores momentos

A continuación calculemos la función generadores de momentos de la variable normal $X \sim N(\mu, \sigma^2)$. En lo que sigue, sea $Z = (x - \mu)/\sigma$ y notemos que $\sigma dz = dx$.

$$\begin{aligned}
M_X(t) &= \mathbb{E}(e^{tX}) \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{x=-\infty}^{\infty} e^{tx} e^{-(x-\mu)^2/(2\sigma^2)} dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{z=-\infty}^{\infty} e^{t\sigma z + t\mu} e^{-z^2/2} \sigma dz \\
&= \frac{1}{\sqrt{2\pi}} e^{\mu t} \int_{z=-\infty}^{\infty} e^{-(z-t\sigma)^2/2 + (t\sigma)^2/2} dz \\
&= \left(e^{t^2\sigma^2/2 + \mu t}\right) \frac{1}{\sqrt{2\pi}} \int_{z=-\infty}^{\infty} e^{-(z-t\sigma)^2/2} dz \\
&= e^{t^2\sigma^2/2 + \mu t}.
\end{aligned}$$

Aquí en la última igualdad usamos el hecho de que:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-t\sigma)^2/2} dz = 1;$$

Para ver esto, notemos que:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(z-t\sigma)^2/2} dz$$

es la función de distribución de una variable aleatoria normal con media $t\sigma$ y desviación estándar 1 y así cuando x converge al infinito la integral es 1.

Podemos usar la función generadora de momentos para verificar nuestro cálculo de la esperanza y la varianza de la distribución normal.

$$M_X^{'}(t) = (\mu + t\sigma^2) e^{t^2\sigma^2/2 + \mu t}$$

y

$$M_X^{''}(t) = (\mu + t\sigma^2)^2 e^{t^2\sigma^2/2 + \mu t} + \sigma^2 e^{t^2\sigma^2/2 + \mu t}.$$

Así $\mathbb{E}(X) = M^{'}(0) = \mu,\ \mathbb{E}(X^2) = M^{''}(0) = \mu^2 + \sigma^2$ y además $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Otra propiedad importante de la distribución normal es que una combinación lineal de variables aleatorias normales tiene una distribución normal:

Teorema 1.1 Sean X e Y variables aleatorias independientes con distribuciones $N(\mu_1, \sigma_1^2)$ y $N(\mu_1, \sigma_2^2)$, respectivamente. Entonces X + Y se distribuye según la distribución normal $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

La distribución normal también puede ser escrita de la siguiente forma:

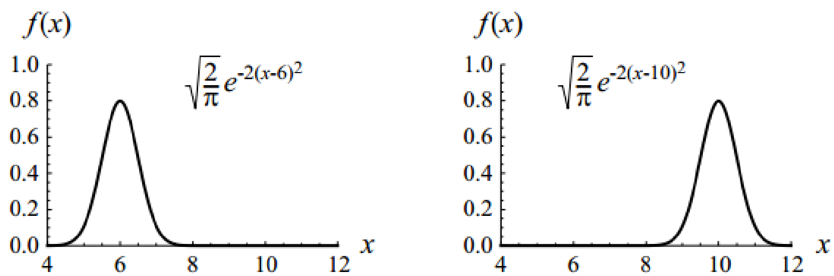
$$f(x) = \sqrt{\frac{b}{\pi}} e^{-b(x-\mu)^2},$$

Puedes indicar explícitamente los parámetros que aparecen, escribiendo la distribución como $f_{\mu,b}(x)$ o $f_{\mu,\sigma}(x)$. La distribución normal es continua y la distribución de probabilidad normal es por lo tanto, una densidad de probabilidad.

Las cantidades μ y b (o μ y σ) afectan la forma y la localización de la curva. Trabajaremos principalmente con la última ecuación que describe la distribución de la variable normal, pero cualquier declaración que hagamos sobre b se puede convertir en enunciados sobre σ reemplazando b con $1/2\sigma^2$.

1.5 Media

Consideramos el siguiente gráfico:



La figura representa con $b = 2$ y $\mu = 6$ y $b = 2$ y $\mu = 10$ dos distribuciones normales:

$$f(x) = \sqrt{\frac{2}{\pi}} e^{-2(x-6)^2} \quad \text{y} \quad f(x) = \sqrt{\frac{2}{\pi}} e^{-2(x-10)^2} \tag{2}$$

De las representaciones se deduce que μ es la ubicación del máximo de la curva. Matemáticamente, esto es cierto porque el factor exponencial $e^{-b(x-\mu)^2}$ tiene un exponente que es cero o negativo (porque un cuadrado siempre es cero o positivo). Entonces, este factor exponencial siempre es menor o igual que 1. Su valor máximo ocurre cuando el exponente es cero, es decir, cuando $x = \mu$.

El pico está por lo tanto ubicado en $x = \mu$. Si aumentamos μ (manteniendo b con el mismo valor), toda la curva simplemente se desplaza hacia la derecha, manteniendo la misma forma.

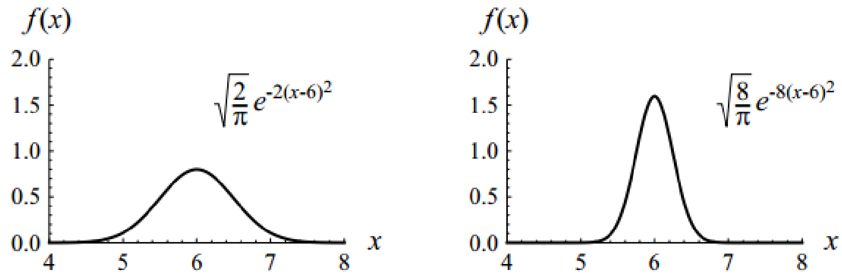
Como la curva es simétrica alrededor del máximo, μ es también la media (o valor esperado) de la distribución: media $= \mu$.

1.6 Altura

Ahora consideremos b. La siguiente figura muestra las gráficas de dos distribuciones normales, una con $b = 2$ y $\mu = 6$ y la otra con $b = 8$ y $\mu = 6$. Las dos funciones son:

$$f(x) = \sqrt{\frac{2}{\pi}} e^{-2(x-6)^2} \quad f(x) = \sqrt{\frac{8}{\pi}} e^{-8(x-16)^2} \tag{3}$$

Debes notar que las escalas en los ejes x e y en esta figura son diferentes a la figuta anterior. La primera función aquí es la misma que la primera función en la figura anterior.



De las representaciones queda claro que b afecta tanto la altura como el ancho de la curva. Veamos cómo se producen estos dos efectos.

El efecto sobre la altura es fácil de entender, porque la altura de la curva (el valor máximo de la función) es simplemente $\sqrt{b/\pi}$. Esto es cierto porque cuando x es igual a μ (que es la ubicación del máximo), el factor $e^{-b(x-\mu)^2}$ es igual a 1, en cuyo caso el valor de $\sqrt{b/\pi}e^{-b(x-\mu)^2}$ es solo $\sqrt{b/\pi}$. (Por el mismo razonamiento, la ecuación (1) da la altura en términos de σ con $1/\sqrt{2\pi\sigma^2}$).

Observando las dos funciones en (3) en observamos que la razón de las alturas es $\sqrt{8/2} = 2$. Y este hecho lo observamos en la última figura.

Para resumir:

$$\text{Altura} = \sqrt{\frac{b}{\pi}} = \sqrt{\frac{1}{2\pi\sigma^2}}.$$

1.7 Ancho en términos de b

Ahora para el ancho. Vemos que la segunda función en la anterior figura es más alta y más estrecha que la primera. (Pero tiene el mismo punto medio, porque no hemos cambiado μ.) El factor por el cual se encoge en la dirección horizontal parece ser aproximadamente $1/2$. Y de hecho, es exactamente $1/2$. Resulta que el ancho de una curva normal es proporcional a $1/\sqrt{b}$. Esto significa que como aumentamos b por un factor de 4 al construir la segunda función, disminuimos el ancho en un factor de $1/\sqrt{4} = 1/2$. Vamos a mostrar ahora que el ancho es de hecho proporcional a $1/\sqrt{b}$.

Pero primero, ¿qué queremos decir con ancho?. Un rectángulo vertical tiene un ancho definido, pero una curva normal no, porque los `lados` están inclinados. Podríamos definir arbitrariamente el ancho de la curva como la altura igual a la mitad de la altura máxima. O en lugar de la mitad, podríamos decir un tercio, o un décimo.

Podemos definirlo como queramos, pero lo bueno es que, el resultado es *proporcional a* $1/\sqrt{b}$ anterior seguirá siendo válido, siempre que elijamos una definición y nos apeguemos a ella para cualquier curva que observemos. Del mismo modo, si queremos trabajar con la ecuación (1) y como $1/\sqrt{b} \propto \sigma$, el ancho será proporcional a σ, independientemente de los detalles de nuestra definición arbitraria.

La definición que elegiremos aquí es: El ancho de una curva es el ancho a la altura igual a $1/e$ (que es aproximadamente 0.37) multiplicado por la altura máxima (que es $\sqrt{b/\pi}$). Esta elección $1/e$ es natural, porque los valores x que corresponden a esta altura son fáciles de encontrar. El resutado es $\mu \pm 1/\sqrt{b}$, ya que la primera expresión en (1) proporciona:

$$\begin{aligned} f(\mu \pm 1/\sqrt{b}) &= \sqrt{b/\pi}e^{-b[(\mu\pm1/\sqrt{b})-\mu]^2} \\ &= \sqrt{b/\pi}e^{-b(\pm1/\sqrt{b})^2} \\ &= \sqrt{b/\pi}e^{-b/b} \\ &= \sqrt{\frac{b}{\pi}} \cdot \frac{1}{e} \end{aligned}$$

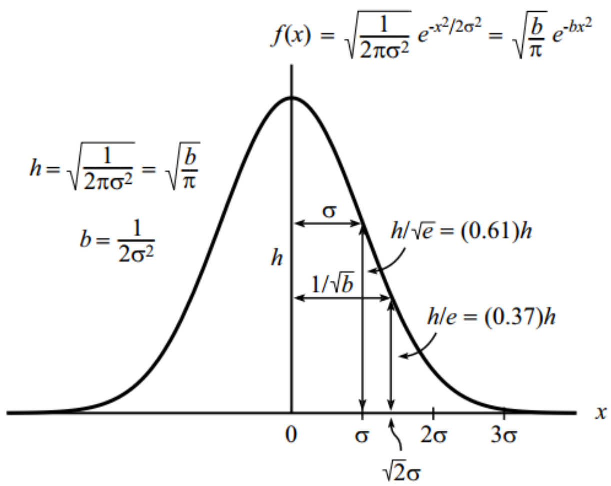
Como la diferencia entre $\mu + 1/\sqrt{b}$ y $\mu - 1/\sqrt{b}$ es igual a $2/\sqrt{b}$, el ancho de la curva normal (según nuestra definición arbitraria) es $2/\sqrt{b}$. Entonces $1/\sqrt{b}$ es la mitad del ancho, que llamaremos 'ancho medio'. (El término 'ancho medio' también puede referirse al ancho total de la curva a la mitad de la altura máxima). Nuevamente, cualquier otra definición de ancho también daría $\sqrt{b}$ en el denominador. Esa es la parte importante. El 2 en el numerador no tiene mucha importancia.

1.8 Ancho en términos de σ

Al trabajar con la ecuación (1), la definición predeterminada del ancho es el ancho en la altura igual a $1/\sqrt{e}$ veces la altura máxima. Esta definición (que es diferente de la definición $1/e$ anterior) se usa porque los valores de x que corresponden a esta altura son simplemente $x \pm \sigma$. Esto es cierto porque si colocamos $x = \mu \pm \sigma$ en la ecuación (1), obtenemos:

$$\begin{aligned} f(\mu \pm \sigma) &= \sqrt{1/2\pi\sigma^2} e^{-[(\mu \pm \sigma) - \mu]^2/2\sigma^2} \\ &= \sqrt{1/2\pi\sigma^2} e^{-(\pm\sigma)^2/2\sigma^2} \\ &= \sqrt{1/2\pi\sigma^2} e^{-1/2} \\ &= \sqrt{\frac{1}{2\pi\sigma^2}} \cdot \frac{1}{\sqrt{e}}. \end{aligned}$$

El factor de $1/\sqrt{e}$ aquí es igual a $1/\sqrt{2.718} \approx 0.61$, que es más grande que el factor $1/e \approx 0.37$ en nuestra definición anterior. Esto es consistente con el hecho de que los puntos $x = \mu \pm \sigma$ (donde la altura es $1/\sqrt{e} \approx 0.61$ veces el máximo) están más cerca del centro que los puntos $x = \mu \pm 1/\sqrt{b} = \mu \pm \sqrt{2}\sigma$ (donde la altura es $1/e \approx 0.37$ veces el máximo). El ancho medio y el resumen de todo lo anterior se muestra en la figura anterior, donde hemos elegido $\mu = 0$ por conveniencia.



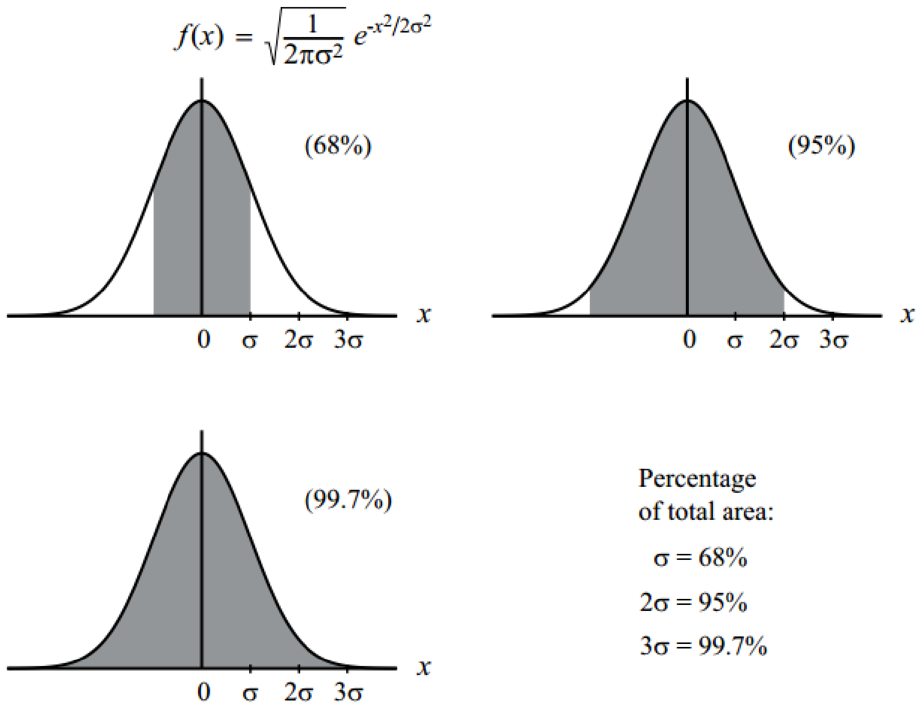
Aunque los puntos $x = \mu \pm \sigma$ producen un buen valor de la distribución normal ($1/\sqrt{e}$ veces el máximo), lo realmente bueno de los puntos $x = \mu \pm \sigma$ es que son una desviación estándar de la media μ. Como se ha demostrado la desviación estándar de la distribución normal dada por la ecuación (1) es simplemente σ. Esta es la razón por la cual la expresión de la ecuación (1) es ampliamente utilizada. Y por la misma razón, la gente generalmente elige (arbitrariamente) definir que el ancho medio de la curva normal sea σ en lugar del ancho medio $1/\sqrt{b} = \sqrt{2}\sigma$ que encontramos anteriormente.

Es decir, están definiendo el ancho mirando dónde la función es $1/\sqrt{e}$ veces el máximo, en lugar de $1/e$ veces el máximo. Como notamos antes, cualquier definición de este tipo está bien; es una cuestión de preferencia personal. El punto crítico es que el ancho es proporcional a σ (o $1/\sqrt{b}$). El factor numérico exacto involucrado es solo una cuestión de definición.

Se puede mostrar numéricamente que alrededor del 68% del área total (probabilidad) bajo la curva normal se encuentra entre los puntos $\mu \pm \sigma$. En otras palabras, se tiene un 68% de posibilidades de obtener un valor de x que esté dentro de una desviación estándar de la media μ.

Usamos la palabra 'numéricamente', porque aunque las áreas debajo de las curvas (o las sumas discretas) para todas las otras distribuciones que se ha mencionado se pueden calcular en forma cerrada, esto no es cierto para la distribución normal. Por lo tanto, cuando encuentres el área bajo la curva normal, siempre deberás especificar los puntos finales numéricos de tu intervalo y luego puede usar una computadora para calcular el área (numéricamente, con la precisión que desee).

También se puede mostrar que el porcentaje del área total que está dentro de dos desviaciones estándar desde μ (es decir, entre los puntos $\mu \pm 2\sigma$) es aproximadamente 95%. Y el porcentaje dentro de tres desviaciones estándar de μ es aproximadamente 99.7%. Estos porcentajes son consistentes con una inspección visual de las áreas sombreadas en la siguiente figura:



Los porcentajes se acercan rápidamente al 100%. El porcentaje dentro de cinco desviaciones estándar de μ es aproximadamente 99.99994%.

Usando la función generadora de momentos también podemos obtener una desviación grande ligada a una variable aleatoria normal.

Teorema 1.2 Sea X una distribución normal $N(\mu, \sigma^2)$, entonces

$$\mathbb{P}\left(\left|\frac{X-\mu}{\sigma}\right| \geq a\right) \leq 2e^{-a^2/2}.$$

Sea $Z = (X-\mu)/\sigma$, entonces por propiedad Z tiene una distribución $N(0,1)$. Luego, para algún $t > 0$,

$$\begin{aligned}\mathbb{P}(Z \geq a) &= \mathbb{P}(e^{tZ} \geq e^{ta}) \\ &\leq \frac{\mathbb{E}(e^{tZ})}{e^{a}} \\ &e^{t^2/2-ta} \\ &e^{-a^2/2},\end{aligned}$$

donde la última desigualdad establecemos $t = a$. El caso $Z \leq a$ de manera similar devuelve el mismo límite, lo que demuestra la afirmación.

2 Propiedades

Hay varias propiedades de simetría importantes que se pueden deducir del PDF y CDF de la distribución normal estándar.

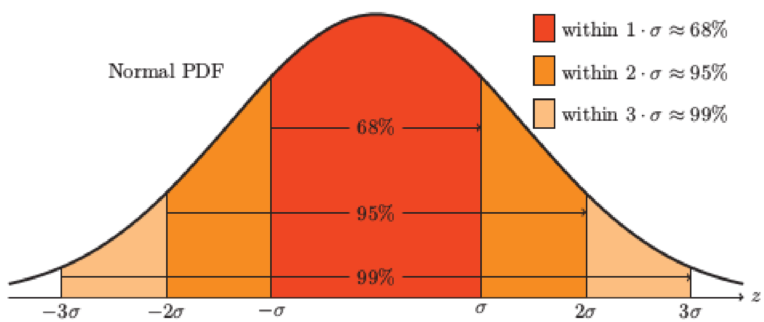
1. La función densidad ϕ es una función par.
2. El área bajo la curva del PDF a la izquierda de -2, que es $P(Z \leq 2) = \Phi(-2)$, por definición, es igual al área a la derecha de 2, que es $\mathbb{P}(Z \geq 2) = 1 - \Phi(2)$. En general, tenemos

$$\Phi(z) = 1 - \Phi(-z)$$

3. Si $Z \sim N(0, 1)$ entonces $-Z \sim N(0, 1)$.

Para hacer aproximaciones es útil recordar la siguiente regla empírica para tres probabilidades aproximadas

$$\mathbb{P}(-1 \leq Z \leq 1) \approx 0.68, \ \mathbb{P}(-2 \leq Z \leq 2) \approx 0.95, \ \mathbb{P}(-3 \leq Z \leq 3) \approx 0.99,$$



Podemos usar la simetría de la distribución normal estándar alrededor de $x = 0$ para hacer algunos cálculos.

Ejemplo 2.1 Calculemos $\Phi(1)$,

Usando el resultado,$\mathbb{P}(-1 \leq Z \leq 1) \approx 0.68$, $\Phi(1) = \mathbb{P}(Z \leq 1)$.

En la figura, las dos colas (en rojo) han combinado el área $1 - 0.68 = 0.32$. Por simetría la cola izquierda tiene área 0.16 (la mitad de 0.32), así que $\mathbb{P}(Z \leq 1) \approx 0.68 + 0.16 = 0.84$.