

# Algunas distribuciones continuas de probabilidad

## 1.1. La distribución normal

**Definición 1.1.** se dice que una variable aleatoria  $X$  se encuentra normalmente distribuida si su función de densidad de probabilidad está dada por

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad \begin{matrix} -\infty < x < \infty \\ -\infty < \mu < \infty, \sigma > 0 \end{matrix}$$

Si se obtienen las dos primeras derivadas de  $f(x; \mu, \sigma)$  con respecto a  $x$  y se igualan a cero, se tiene que el valor máximo de  $f(x; \mu, \sigma)$  ocurre cuando  $x = \mu$ , y los valores  $x = \mu \pm \sigma$  son las abscisas de los dos puntos de inflexión de la curva.

**Demostrar que la definición 5.1 es una función de densidad de probabilidad.**

Demostración.- El que la función sea no negativa se satisface, ya que  $f(x; \mu, \sigma) > 0$  para  $-\infty < x < \infty$ ,  $-\infty < \mu < \infty$  y  $\sigma > 0$ . Para demostrar que:

$$\int_{-\infty}^{\infty} f(x; \mu, \sigma) dx = 1.$$

Sea

$$I = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

el valor de la integral y aplíquese la transformación lineal  $y = (x - \mu)/\sigma$  de manera tal que  $x = \sigma y + \mu$  y  $dx = \sigma dy$ . Esto da como resultado:

$$I = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy.$$

Si puede demostrarse que  $I^2 = 1$ , puede deducirse que  $I = 1$  puesto que  $f(x; \mu, \sigma)$  tiene un valor positivo. De acuerdo con lo anterior:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{(y^2+z^2)}{2}} dydz,$$

en donde se ha escrito el producto de las dos integrales como una doble integral ya que las funciones de  $z$  son contantes con respecto a  $y$  como también de manera viceversa. Al cambiar de coordenadas rectangulares representadas por  $x$  e  $y$ , a coordenadas polares  $r$  y  $\theta$ , en donde  $y = r \cos \theta$  y  $z = r \sin \theta$ . Esto

es:

$$y^2 + z^2 = r^2 \cos^2 \theta + r^2 \sin^2 \theta = r^2$$

y el elemento de área  $dydz$ , en coordenadas rectangulares se reemplaza por  $rdrd\theta$  en coordenadas polares. Dado que los límites  $(-\infty, \infty)$  tanto para  $y$  como para  $z$  generan el plano completo  $yz$ , el plano correspondiente a  $r$  y a  $\theta$  se genera mediante el empleo de los límites  $(0, 2\pi)$  para  $\theta$  y  $(0, \infty)$  para  $r$ . De esta forma se tiene:

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^\infty e^{-r^2/2} r dr d\theta = \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_0^\infty e^{-r^2/2} r dr = \frac{\theta}{2\pi} \Big|_0^{2\pi} \cdot [-e^{-r^2/2}] \Big|_0^\infty = 1.$$

**La media de una variable aleatoria distribuida normalmente** se encuentra definida por:

$$E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Se pretende demostrar que  $E(X) = \mu$ . Supóngase que a  $E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$  se suma y se resta

$$\frac{\mu}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

La identidad se mantiene, pero después de reacomodar términos se tiene

$$\begin{aligned} E(X) &= \int_{-\infty}^\infty (x - \mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \frac{\mu}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x (x - \mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \mu \end{aligned}$$

dado que el valor de la segunda integral es uno. Al afectar un cambio de variable de integración de manera tal que  $y = \frac{x - \mu}{\sigma}$ ,  $x = \sigma y + \mu$  y  $dx = \sigma dy$ , se tiene:

$$E(X) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^\infty y e^{-y^2/2} dy + \mu = -\frac{\sigma}{\sqrt{2\pi}} e^{-y^2/2} \Big|_{-\infty}^\infty + \mu = \mu.$$

El lector recordará de sus cursos de cálculo que la última integral es cero porque el integrando es una función impar y la integración se lleva a cabo sobre un intervalo simétrico alrededor de cero.

Si el valor máximo de la función de densidad de probabilidad normal ocurre cuando  $x = \mu$  este es la media, la mediana y la moda de cualquier variable aleatoria distribuida aleatoriamente.

Para encontrar los demás momentos, se determinará la función generadora de momentos. Por definición:

$$m_{X-\mu}(t) = E[e^{t(X-\mu)}] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{t(x-\mu)} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{-\frac{1}{2\sigma^2}[(x-\mu)^2 - 2\sigma^2 t(x-\mu)]} dx.$$

de donde se completa el cuadrado en el interior del paréntesis rectangular y se tiene:

$$(x - \mu)^2 - 2\sigma^2 t(x - \mu) = (x - \mu)^2 - 2\sigma^2 t(x - \mu) + \sigma^4 t^2 - \sigma^4 t^2 = (x - \mu - \sigma^2 t)^2 - \sigma^4 t^2.$$

Por lo que,

$$m_{X-\mu}(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{\frac{\sigma^2 t^2}{2}} e^{-\frac{(x-\mu-\sigma^2 t)^2}{2\sigma^2}} dx = e^{\frac{\sigma^2 t^2}{2}},$$

dado que el integrando junto con el factor  $\frac{1}{\sqrt{2\pi}\sigma}$  es una función de densidad de probabilidad normal con parámetros  $\mu + \sigma^2 t$  y  $\sigma$ .

Al desarrollar  $e^{\frac{\sigma^2 t^2}{2}}$  en serie de potencias se tiene:

$$m_{X-\mu}(t) = 1 + \frac{(\sigma t)^2}{2} + \frac{(\sigma t)^4}{4 \cdot 2!} + \frac{(\sigma t)^6}{8 \cdot 3!} + \frac{(\sigma t)^8}{16 \cdot 4!} + \dots$$

Cuando las potencias impares de  $t$  no se encuentran presentes, todos los momentos centrales de  $X$  de orden impar son cero, de esta forma se asegura la simetría de la curva.

La segunda derivada de  $m_{X-\mu}(t)$  evaluada en  $t = 0$  es **la varianza** y está dada por:

$$Var(X) = \left. \frac{d^2 m_{X-\mu}(t)}{dt^2} \right|_{t=0} = \sigma^2 + \frac{12t^2\sigma^4}{4 \cdot 2!} + \frac{30t^4\sigma^6}{8 \cdot 3!} + \dots \Big|_{t=0} = \sigma^2;$$

De esta manera **la desviación estándar es  $\sigma$** . De manera similar, la cuarta derivada de  $m_{X-\mu}(t)$  evaluada en  $t = 0$  es el cuarto momento central, el cual es:

$$\mu_4 = \left. \frac{d^4 m_{X-\mu}(t)}{dt^4} \right|_{t=0} = 3\sigma^4 + \frac{360t^2\sigma^6}{8 \cdot 3!} + \dots \Big|_{t=0} = 3\sigma^4$$

De acuerdo con lo anterior, para cualquier distribución normal el coeficiente de asimetría es  $\alpha_3(X) = 0$ , mientras que la curtosis relativa es  $\alpha_4(X) = \frac{3\sigma^4}{\sigma^4} = 3$ . Para momentos alrededor del cero, puede determinarse la función generadora de momentos centrales o viceversa. Dado que

$$m_{X-\mu}(t) = E \left[ e^{t(X-\mu)} \right] = e^{-\mu t} E \left[ e^{tX} \right] = e^{-\mu t} m_X(t),$$

para una distribución normal

$$e^{-\mu t} m_X(t) = e^{\frac{\sigma^2 t^2}{2}}$$

y

$$m_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

La probabilidad de que una variable aleatoria normalmente distribuida  $X$  sea menor o igual a un valor específico,  $x$  está dada por **función de distribución acumulativa**

$$P(X \leq x) = F(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt.$$

Sea  $Z$  una variable aleatoria definida por la siguiente relación:

$$Z = \frac{(X - \mu)}{\sigma}$$

en donde  $\mu$  y  $\sigma$  son la media y la desviación estándar de  $X$ , respectivamente. De acuerdo con lo anterior,  $Z$  es una variable aleatoria estandarizada con media cero y desviación estándar uno. Así,

$$P(X \leq x) = P[X \leq (x - \mu)/\sigma] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{(x-\mu)/\sigma} e^{\frac{-z^2}{2}} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{\frac{-z^2}{2}} dz.$$

El integrando junto con el factor  $1/\sqrt{2\pi}$  es la **función de densidad de probabilidad de la variable aleatoria normal estandarizada Z**. De donde

$$F_X(x; \mu, \sigma) = F_Z(z; 0, 1)$$

Para cualquier valor específico de  $z$ , el correspondiente valor en la tabla es la probabilidad de que la variable aleatoria normal estándar  $Z$  sea menor o igual a  $z$ ; esto es

$$P(Z \leq z) = F_Z(z; 0, 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt.$$

La notación  $X \sim N(\mu, \sigma)$  denotará que la variable  $X$  se encuentra distribuida normalmente con media  $\mu$  y desviación estándar  $\sigma$ .

Determinaremos la probabilidad de que un valor de  $X$  se encuentre entre  $a$  y  $b$ , si  $X \sim N(\mu, \sigma)$ . Por definición:

$$P(a \leq X \leq b) = \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx,$$

pero, mediante el empleo de  $E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x (x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \mu$  se tiene:

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{z^2}{2}} dz = F_Z\left(\frac{b-\mu}{\sigma}; 0, 1\right) - F_Z\left(\frac{a-\mu}{\sigma}; 0, 1\right)$$

**Teorema 1.1.** Sea  $X$  una variable aleatoria binomial con media  $np$  y desviación estándar  $\sqrt{np(1-p)}$ . La distribución de la variable aleatoria tiende a la normal

$$Y = \frac{X - np}{\sqrt{np(1-p)}}$$

estándar conforme el número de ensayos independientes  $n \rightarrow \infty$ .

**Demostración.-** La demostración que aquí se presenta se basa en el hecho de que una función generadora de momentos define, de manera única, a una distribución. Se demostrará que la función generadora de momentos de  $Y$  tiende a una distribución normal conforme  $n \rightarrow \infty$ .  $X$  es una variable aleatoria binomial:

$$m_X(t) = [(1-p) + pe^t]^n$$

Entonces:

$$m_Y(t) = E(e^{tY}) = E\left[e^{\frac{t(X-np)}{\sqrt{np(1-p)}}}\right] = e^{\frac{npt}{\sqrt{np(1-p)}}} \cdot E\left[e^{\frac{tX}{\sqrt{np(1-p)}}}\right]$$

donde  $E\left[e^{\frac{tX}{\sqrt{np(1-p)}}}\right]$  es la función generadora de momentos de  $X$  con argumento  $\frac{t}{\sqrt{np(1-p)}}$ . De esta forma se tiene:

$$m_Y(t) = e^{\frac{-npt}{\sqrt{np(1-p)}}} \left[(1-p) + pe^{\frac{t}{\sqrt{np(1-p)}}}\right]^n;$$

pero: