單元 27: 常態機率分布
(課本 §4.5)

最被廣泛使用的連續機率分布為常態分布, 其 pdf 呈現鐘型, 正式定義如下.

定義 1. $Y \sim N(\mu, \sigma)$, $-\infty < \mu < \infty$, $\sigma > 0$ 若且為若機率密度函數 pdf

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, -\infty < y < \infty$$

並稱隨機變數 Y 有常態機率分布 (normal probability distribution).

註. 常態隨機變數的 pdf 有兩個參數 μ 與 σ.

定理 4.7. 設 $Y \sim N(\mu, \sigma)$, 則期望值

$$E(Y) = \mu$$

且變異數

$$\text{Var}(Y) = \sigma^2$$

<證> (略) 延至 §4.9.

註 1. 機率密度函數

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)},\ -\infty < y < \infty$$

的圖形如下.

最大值爲 $\frac{1}{\sigma\sqrt{2\pi}}$; 兩個反曲點發生在距 μ 左右各 σ 長度的位置.

註 2. 隨機變數 Y 落在 $[a, b]$ 區間內的機率

$$P(a \le Y \le b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)} dy$$

的完整型 (closed form, 眞確值) 不存在, 故需以附錄 III 的表 4 求其近似值. 表 4 僅列出向右偏離期望值 μ 有 z 個標準差以外的區域面積的近似值 $A(z)$ (稱作右尾機率, right-tail probability), 如下圖所示. 但透過對稱性可求一般區域的面積, 如 $P(a \le Y \le b)$ 的近似值, 說明如下例.

例 1. 令 $Z \sim N(0, 1)$ (稱作標準常態分布). 則

(a) 因爲 $\mu = 0$, $\sigma = 1$, 故 $Z > 2 \Leftrightarrow$ 由 μ 向右偏離 2 個標準差 (std) 以外. 因此, 根據 $A(z)$ 的定義及表 4,

$$P(Z > 2) = A(2) = 0.0228$$

(b) 如圖所示,

$$
\begin{aligned}
& P(-2 \leq Z \leq 2) \\
= \; & \text{在 } [-2,2] \text{ 上 pdf 所圍出的面積} \\
= \; & 1 - \text{pdf 在 } -2 \text{ 左邊及 } 2 \text{ 右邊所圍出的面積} \\
= \; & 1 - A(2) - A(2) \text{ (根據對稱性)} \\
= \; & 1 - 2(0.0228) \\
= \; & 1 - 0.0456 = 0.9544
\end{aligned}
$$

(c) 如圖所示,

$$
\begin{aligned}
& P(0 \leq Z \leq 1.73) \\
= \; & \text{pdf 在 } [0, 1.73] \text{ 上所圍出的面積} \\
= \; & \text{縱軸右邊去掉 } 1.73 \text{ 右邊所圍出的面積} \\
= \; & 0.5 - A(1.73) \\
= \; & 0.5 - 0.0418 = 0.04582
\end{aligned}
$$

例 2. 設入學考試的成績分布 $\sim N(75, 10)$, 則如圖所示,

$$
\begin{aligned}
& P(80 \leq \text{成績} \leq 90) \\
= \; & \text{pdf 在 } [80, 90] \text{ 上所圍出的面積} \\
= \; & \text{pdf 在 } 80 \text{ 右邊所圍出的面積 } - \\
& \qquad \text{在 } 90 \text{ 右邊所圍出的面積} \qquad (1)
\end{aligned}
$$

又 80 的右邊 $\Leftrightarrow$ 由 $\mu = 75$ 向右偏移 $\frac{80-75}{10}$ 個標準差以外. 同理, 90 的右邊 $\Leftrightarrow$ 向右偏移 $\frac{90-75}{10}$ 個標準差以外. 因此, 接續 (1) 式並根據表 4,

$$
\begin{aligned}
&P(80 \le \text{成績} \le 90) \\
&\quad = A\left(\frac{80-75}{10}\right) - A\left(\frac{90-75}{10}\right) \\
&\quad = A(0.5) - A(1.5) \\
&\quad = 0.3805 - 0.0668 = 0.2417
\end{aligned}
$$

註. 若 $Y \sim N(\mu, \sigma)$ 則

$$
Z \overset{\text{def}}{=} \frac{Y-\mu}{\sigma} \sim N(0, 1)
$$

證明 (略), 延至第六章.