

## 第五章 常態曲線 (The Normal Curve)

### 壹、本單元的目標

- 1、定義並說明常態曲線 (the normal curve) 的概念。
- 2、學習將資料原來得到的數值轉換成 Z 分數 (z scores)，以及運用 Z 分數及常態曲線表 (見 Appendix A) 來找出在曲線上某一點之上、之下或某兩點之間的面積。
- 3、以機率的方式來表達前述之面積。

### 貳、前言

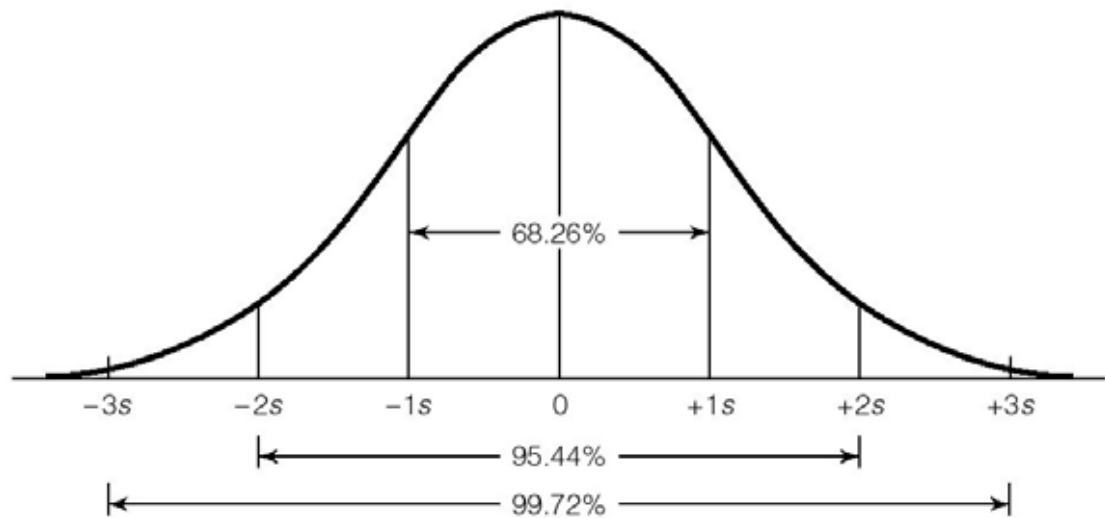
常態曲線 (The Normal Curve) 及常態分配的觀念，在統計中十分重要，它們是推論統計的基礎。

常態曲線及分配是一種理論模式，但透過這理論模式，配合平均數及標準差，我們可以對實證研究所得之資料分配，做相當精確之描述及推論。能做到這一點是因常態曲線本身有些重要且已知的特性。常態曲線最重要的特性是其形狀為左右對稱若鐘形之曲線。此曲線只有一個眾數，並與中位數及平均數是三合一的。其區線的兩尾是向兩端無限延伸。因此，雖然實際調查得到的資料，不可能是這種完美的理論模式，但許多實際得到之變項的資料分配是相當接近這種模式，因此可以假定它們的分配是常態的，進而使我們得以運用常態曲線的理論特性。

配合平均數及標準差之觀念，我們可以得到常態分配一個重要的特性：在常態曲線下，以平均數  $\bar{X}$  為中心，任何一個在左邊的點與  $\bar{X}$  之間在常態曲線下之面積是和另一相對在右邊同距離之點與  $\bar{X}$  之間的面積相等。

常態分配另一非常重要特性是，任何點與  $\bar{X}$  間在常態曲線下之面積是一定且已知的。見〈圖一〉

#### AREAS UNDER THE THEORETICAL NORMAL CURVE



圖一、在常態曲線下之面積

從〈圖一〉可知在常態曲線下，平均數與標準差之間的所佔的面積比例是有一定的關係。您要熟記這個關係。如果一個變項的分配是接近常態曲線，那這個面積的比例也代表所佔的樣本比例。例如，如果全部樣本數是 1000 人，則平均數加減一個標準差（平均數  $\pm 1S$ ）就有約 683 人（ $1000 \times 68.26\%$ ）。所以，就常態分配而言，只有少數的樣本是在平均數加減三個標準差以外（也就是說只有極少數個案的分數是比平均數加三個標準差來的大，或比平均數減三個標準差來的小）。

#### 參、標準化：求 Z scores

在我們了解以上常態分配之特性後，可以進一步將實際之分配加以標準化成為標準常態分配（**standard normal distribution**）。這種分配之特性是其  $\bar{X} = 0$ ， $S = 1$ 。（※有兩點要注意：1、實際分配應是接近常態分配，標準化才有意義。但如果實際分配不是接近常態分配時，常態分配之觀念仍然有用，容以後再述。2、常態分配如前所言是一種理論模式，也就是一種用來了解現實狀況的標準）。

資料標準化方法是將原來資料中的分數變成 Z scores（Z 分數），一種標準常態分配之分數。原來的分數可以是任何單位測量到的，如「元」、「歲」或「分」。在轉變成 Z 分數後，這些單位就消失了，而原來的平均數會成為 0，原來的標準差則成為 1。例如，在經過智力測驗後，志明的 IQ 分數是 120 分，而此分數是比整個樣本的平均數多一個標準差，也就是 10 分。當整個樣本的 IQ 的分數轉換成 Z

分數後，整個樣本的平均數是 0，而志明的 IQ 分數也就成了 1。

轉換原始分數成為 Z 分數的公式為：

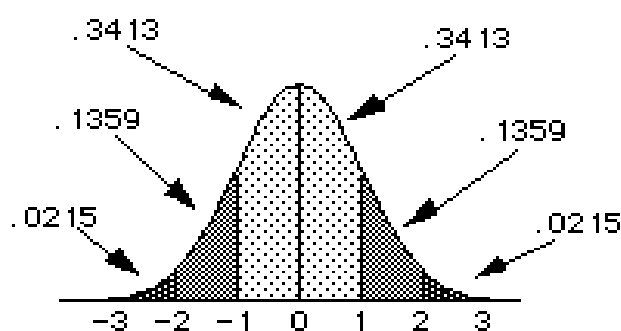
$$Z = \frac{X_i - \bar{X}}{S}$$

由此公式可知，當  $X_i = \bar{X}$  時， $Z=0$ ，也就是  $\bar{X}$  在標準常態分配下的 Z 分數等於 0。此外，一個原來分數等於原來的  $\bar{X}$  加上一個 S 時，經由公式(1)之轉換，即成  $Z=1$ ，即  $X_i = \bar{X} + 1S$  時，

$$Z = \frac{(\bar{X} + 1S) - \bar{X}}{S} = 1$$

這也和剛才提及在標準常態分配中  $S=1$  之說。同理可知，當一個原始分數轉成 Z 分數是 1 時，那此原始分數就是比平均數高一個標準差的分數。

標準化之觀念即將原來之分數化成一種標準分數，如同將英尺化成公尺般。如此一來，不同之樣本分配經標準化後就可比較。而原來的分數所構成的常態分配也就成了標準常態分配。當然，標準常態分配也有前述常態分配所有的特性，如下圖：



圖二、標準常態曲線下之面積

#### 肆、常態曲線表

任何分數經標準化成 Z 分數後，此 Z 分數與  $\bar{X} (=0)$  之間在常態曲線下之面積，可利用書中 Appendix A 之常態曲線表查出。此表有三欄，其中一部分如下表：

**TABLE 5.1** AN ILLUSTRATION OF HOW TO FIND AREAS UNDER THE NORMAL CURVE USING APPENDIX A

(a) Z	(b) Area Between Mean and Z	(c) Area Beyond Z
0.00	0.0000	0.5000
0.01	0.0040	0.4960
0.02	0.0080	0.4920
0.03	0.0120	0.4880
⋮	⋮	⋮
1.00	0.3413	0.1587
1.01	0.3438	0.1562
1.02	0.3461	0.1539
1.03	0.3485	0.1515
⋮	⋮	⋮
1.50	0.4332	0.0668
1.51	0.4345	0.0655
1.52	0.4357	0.0643
1.53	0.4370	0.0630
⋮	⋮	⋮

© 2005 Wadsworth - Thomson

利用此表，我們可知道： 1、任何 Z 分數和  $\bar{X}$  之間的面積； 2、某 Z 分數以上或以下之面積（也就是找大於或小於 Z 分數之部分的面積）； 3、兩個 Z 分數間之面積。

在常態曲線下，尋找某 Z 分數以上或以下之面積的方法，如下表：

表一、尋找某 Z 分數以上或以下之面積的方法（對照 Appendix A）

尋找面積	當 Z 分數為	
	正	負
在 Z 分數以上（比 Z 分數大的部分）	看 c 欄	將 b 欄所呈現的面積加 .5000 或 50%
在 Z 分數以下（比 Z 分數小的部分）	將 b 欄所呈現的面積加 .5000 或 50%	看 c 欄

在常態曲線下，尋找兩個 Z 分數之間面積的方法如下表：

表二、尋找兩個 Z 分數之間面積的方法（對照 Appendix A）

狀況	尋找步驟
兩個 Z 分數都在平均數的同一邊	由 b 欄找到每一個 Z 分數與平均數之間的面積，然後將大的面積減去小的面積
兩個 Z 分數是分在平均數的兩邊	由 b 欄找到每一個 Z 分數與平均數之間的面積，然後將兩個面積相加

#### 伍、用常態曲線來推估機率(probability)

有已知理論特性的常態曲線也可看成是機率的分配。透過我們對常態分配之特性的了解，我們可以進而了解資料中任何一件個案或分數，在一定條件下被選出或抽中之機率。找到機率的方法就與上述找面積的方法相同。

所謂機率，簡言之，即從長期之觀點來看，某一事件發生或成功(success)與全部事件間之比例關係。比如說，從 52 張撲克牌中抽取紅心 Q 之比例，即機率應是  $1/52$ 。當然您要是只做一次抽牌實驗，抽出的極可能不是紅心 Q，但是要是做一百次或一千次，則平均下來，抽到紅心 Q 之比例應是接近  $1/52$ ，這也就是所謂的「從長期之觀點來看(over the long run)」的意思。機率一定是在 0 與 1 之間。機率是 0 時，就表示某一事件毫無發生的可能，而機率為 1 時，則表示此事件必然會發生。

當一個變項的分配是常態分配時，我們就可利用常態分配的特性來估計抽中某些樣本是有某些分數的機率為何。例如，如果男性的 IQ 分數是常態分配，則我們可以很容易知道抽中男性的 IQ 分數是在 95 分與平均數 100 分間的機率為何。其作法就如先前將原始分數(95 分與 100 分)先轉換成 Z 分數後，再尋找此二分數間的面積。因為，所謂 68% 左右之面積是在  $\bar{X} \pm 1S$  之間的意義，即為有 68% 左右之個案件數或分數是在  $\bar{X} \pm 1S$  之間。也因此，從機率的觀點來看，當我們隨便由一有常態分配性質之母體抽一件個案時，非常可能的是此件之分數是在  $\bar{X} \pm 1S$  之間，由此可推，任何一件超過  $\bar{X} \pm 3S$  之分數並抽到的可能性極低應只有 0.0028 之概率( $1 - 0.9972$ )。

#### 陸、分配之偏態(Skewness)與峰度(Kurtosis)

$$\text{Skewness} = \frac{\sum (X_i - \bar{X})^3 / N}{S^2 \times S} = \frac{\sum (X_i - \bar{X})^3}{NS^3} = \frac{\sum Z^3}{N} = \overline{Z^3}$$

$$\text{Kurtosis} = \frac{\sum (X_i - \bar{X})^4 / N}{(S^2)^2} = \frac{\sum (X_i - \bar{X})^4}{NS^4} = \frac{\sum Z^4}{N} = \overline{Z^4}$$

常態分配之 Skewness 為 0，Kurtosis 為 0。