



# 產銷資訊系統

## 預測

李曉惠 (產) & 李怡慧 (銷／系統)



# 聽到預測你想到甚麼？



# 甚麼叫做好的預測？



# 需要預測的例子？

- 如7-11的鮭魚御飯糰的每日需求
- 還有嗎？



# 預測的兩種用途

- 系統規劃

- 包含一些長程規劃，例如，提供何種型態的產品與服務、籌備設備與機具、地點設置等。

- 系統使用的規劃

- 指一些中短程規劃，例如，存貨規劃、勞動力水準、採購與生產規劃、預算與排程規劃等。



# 預測需求

- 企業預測不只針對需求做預測，但為了簡化討論，本章的焦點只放在預測需求。謹記，此概念與方法也能夠用在其他的預測變數上。
- 不管是使用電腦或複雜的數學模型，預測都不能算是精確的科學。
- 要發展成功的預測，經驗、判斷能力與專業技術扮演著重要的角色。



# 預測技術

- 常用的預測技術，分為定性法與定量法。
  - 定性法
    - 主要由主觀判斷所組成，通常缺乏精確的數字描述。
    - 能在預測流程中加入俗稱軟性的資訊（例如，人的因素、個人意見、直覺）。
  - 定量法
    - 涵蓋客觀的歷史資料，或開發以因果變數做預測的關聯性模型。
    - 由客觀分析或硬性資料所組成，通常可以避免會影響定性法結果的個人偏見。



# 各種預測技術的共同特徵

- 預測的技術的共同特徵為：
  1. 預測技術通常假設過去存在的因果系統未來將持續。
  2. 預測很少 **完美無缺**。
  3. 整體預測會比單一預測更為精確。
  4. 隨著預測的時間週期愈廣，亦即增加**時間幅度 (time horizon)**，預測精確度會減少。
  5. 彈性的（亦即能迅速反應需求改變）企業組織進行預測所需的幅度較短。





# 三種預測技術

- **判斷預測法(judgmental forecast)**
  - 根據主觀的分析。
- **時間序列預測法(time-series forecast)**
  - 根據過去的經驗來了解未來，利用歷史資料，並假設未來和過去所有情況相同。
- **關聯性模型(associative model)**
  - 利用包含一個以上可解釋的變數方程式來預測需求。



# 定性預測

## ● 主管的意見

- 高層管理者一同進行預測。

優點	缺點
集合各個部門管理者重要的知識與智慧。	可能某位管理者的意見勝出，整個團體便傾向此看法，無法綜合意見。

## ● 銷售員的意見

- 以銷售／客服人員為資訊來源。

優點	缺點
銷售人員了解顧客的需求，並能考慮到未來性。	銷售人員難以分辨顧客想做的事與實際會做的事。其亦可能會受到最近銷售經驗的影響。若預測被用來定銷售配額，容易導致利害衝突。



# 定性預測

## ● 消費者調查

- 組織設計問卷，並以抽樣方式直接從顧客身上獲得資訊。

優點	缺點
可以得到別種方法得不到的資訊。	<ol style="list-style-type: none"><li>1. 需相當程度的知識和技能才能進行調查、更正與解釋錯誤資訊結果。</li><li>2. 調查相當昂貴且耗時。</li><li>3. 需考慮不合理的行為模式。</li></ol>

## ● 德菲法(Delphi method)

- 反覆進流程以取得一致的預測結果。



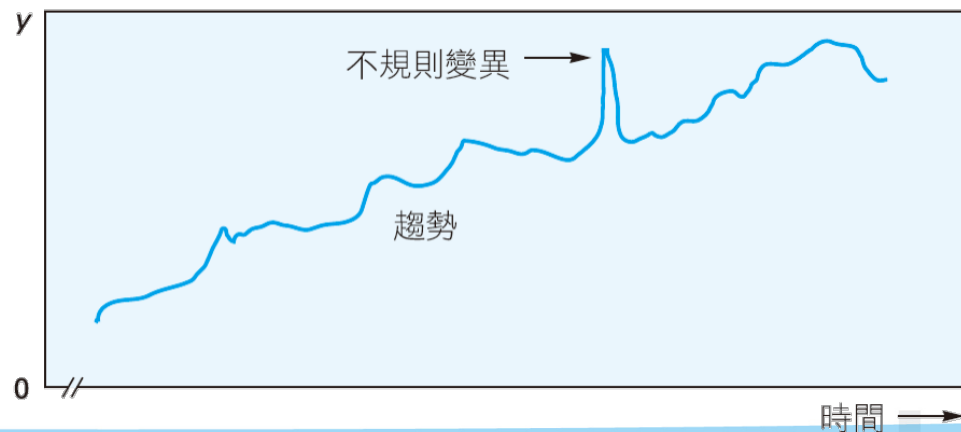
# 以時間序列資料為基礎的預測

- 時間序列(time series)

- 指間隔固定時間並依時間順序排列的觀測值。
- 假設未來的序列值能經由過去的序列值估計。

- 時間序列的基本行為：

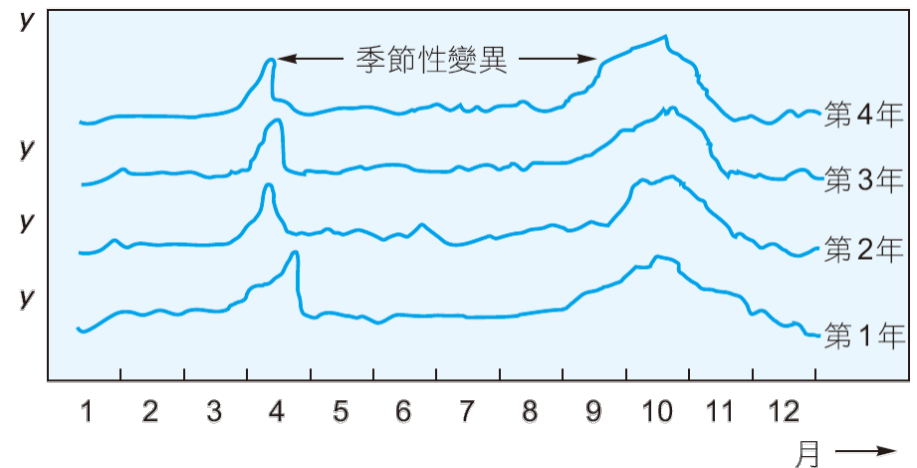
- 趨勢(trend)：指資料長期隨時間向上或向下的移動。



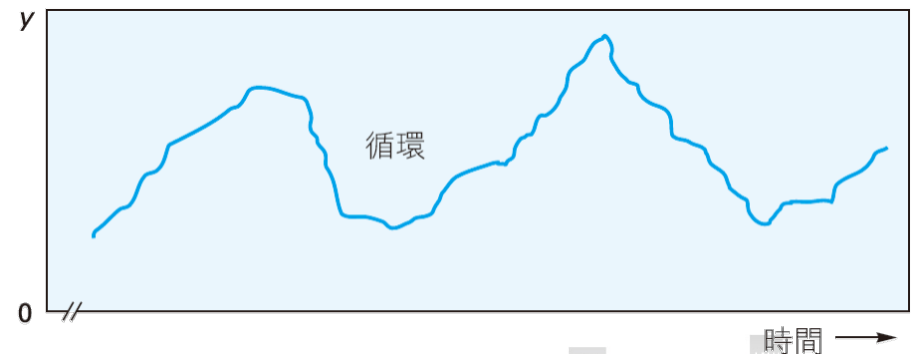


# 時間序列基本行為

- **季節性(seasonality)**：指短期、規則的變異，通常與日期或一天中的時間因素有關。



- **循環(cycle)**：指持續一年以上的波狀變異。





# 時間序列基本行為

- **不規則變異(irregular variation)**：指由不尋常的情況所產生，其無法反映典型行為，應盡可能確認出這些資料並移除。
- **隨機變異(random variations)**：指考量所有其他活動狀況之後，所留下的殘餘變異。



# 天真預測法(Naive Forecast)

- 使用時間序列的前一期數值當做預測基礎，可用於穩定序列、季節性變異或趨勢。

優點	缺點
不需任何成本、方法簡單迅速； 無需分析資料，故容易了解。	不能提供高精確度的預測。

穩定序列	上一期的數值將成為下一期的預測值。
季節性變異	本季的預測值和上一季的數值相同。
趨勢	預測值為序列中，前一個數值加上（或減去）最後兩個數值的差。



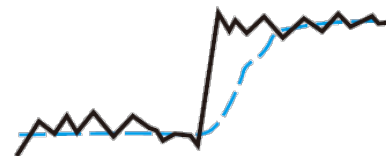
# 平均法分析技術

- 歷史資料通常包含部分的隨機變異或白噪音(**white noise**)，平均法分析技術可以使得資料的變異變小

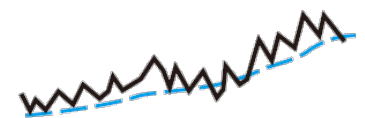
- 適合此法的三種型態



理想型態



階梯型變動型態  
(預測落後)



漸進型變動型態  
(預測落後)

- 三種平均法技術：

- 移動平均法。
- 加權移動平均法。
- 指數平滑法。





# 移動平均法(Moving Average)

- 使用數個近期的實際資料來產生預測值。

優點	缺點
容易計算與了解。	各期的資料權重都相同。

$$F_t = MA_n = \frac{\sum_{i=1}^n A_{t-i}}{n} = \frac{A_{t-n} + \cdots + A_{t-2} + A_{t-1}}{n}$$

其中

$F_t$  = 第  $t$  期的預測值

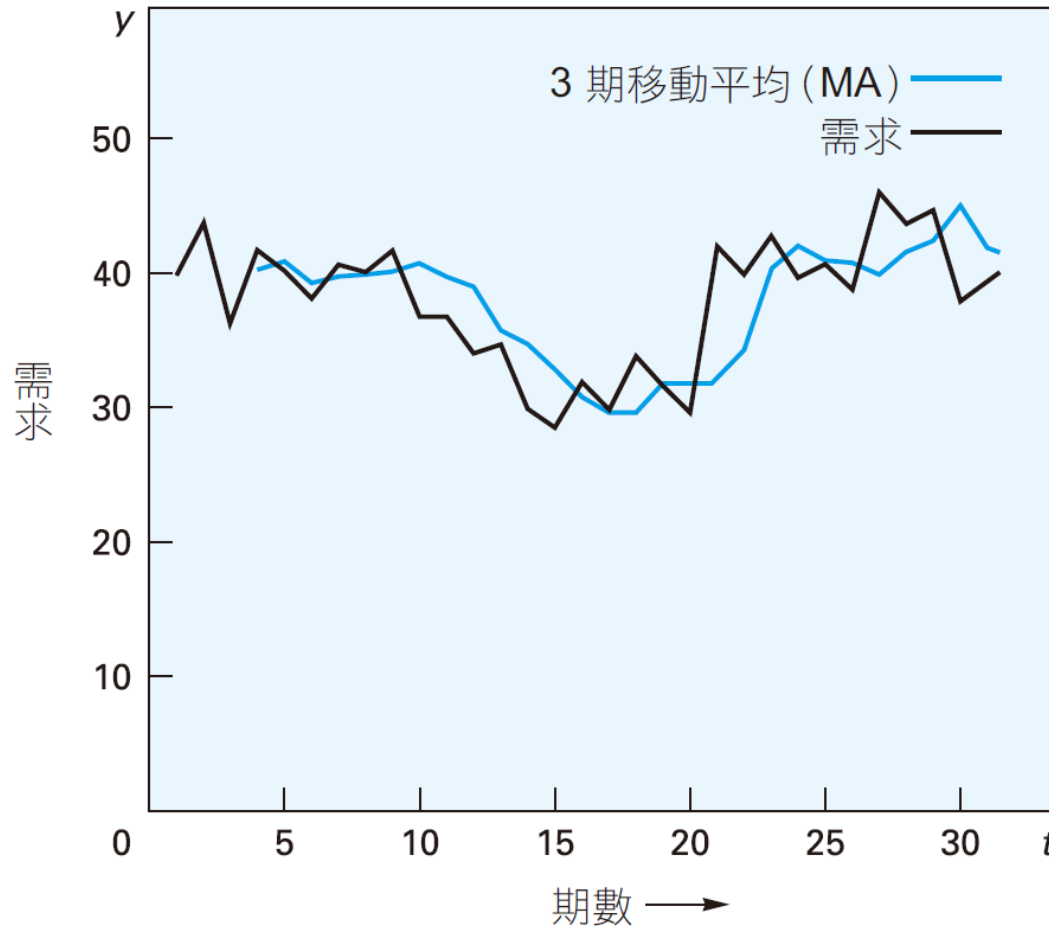
$MA_n$  =  $n$  期的移動平均

$A_{t-i}$  = 第  $t - i$  期的實際值

$n$  = 移動平均的期數 (資料點)



# 移動平均法(Moving Average)



說明：  
移動平均法的資料傾向於平滑，而且變動較實際資料落後。



# 麵包店練習

- 國際掃描公司專賣倉庫標籤掃描器，7個月來的銷售量如右
- 使用天真預測法，預測9月的銷售量
- 用3個月的移動平均法？
- 用5個月的移動平均法？

月	銷售量（單位為千台）
2月	19
3月	18
4月	15
5月	20
6月	18
7月	22
8月	20





# 加權平均法(Weighted Average)

- 和移動平均法很類似，不同之處是愈近期的資料，給定的權重愈大（權重總和必為1.00）。

優點	缺點
與移動平均法相比，其能反映較近期的事情。	權重的選擇有點武斷，通常需使用試誤法才能找出適當的權重。

$$F_t = w_t(A_t) + w_{t-1}(A_{t-1}) + \cdots + w_{t-n}(A_{t-n})$$

其中  $W_t$  = 期間  $t$  之權重， $W_{t-1}$  = 期間  $t - 1$  之權重  
 $A_t$  = 期間  $t$  實際值， $A_{t-1}$  = 期間  $t - 1$  之實際值



# 指數平滑法(Exponential Smoothing)

- 每一個新預測值以前一個預測值為基礎，再加上預測值與實際值差額的百分比。

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

其中

$F_t$  = 第 $t$ 期的預測值

$F_{t-1}$  = 前一期的預測值（亦即第 $t-1$ 期）

$\alpha$  = 平滑常數

$A_{t-1}$  = 前一期的實際需求或銷售量

上述公式亦可寫作：

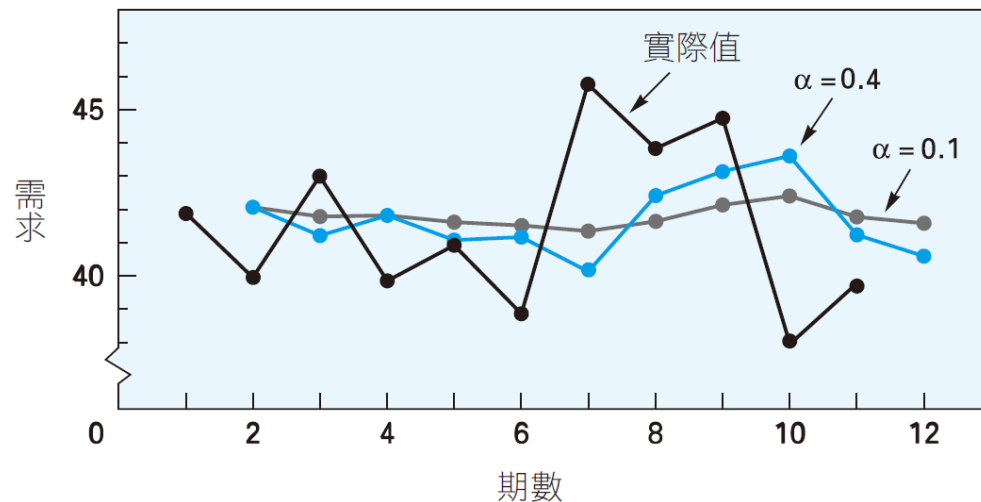
$$F_t = (1 - \alpha)F_{t-1} + \alpha A_{t-1}$$



# 指數平滑法(Exponential Smoothing)

- 平滑常數 $\alpha$ 愈接近0，則預測誤差調整的速度愈平滑。

優點	缺點
計算容易，僅需改變 $\alpha$ 即可變更權重。	為一複雜的加權平均法。





# 回到麵包店練習

- 如果使用三個月加權平均法，前一月權重為0.60，前兩月權重為0.30，前三月權重為0.10
- 如果使用指數平滑法呢？  
假設平滑常數為0.20，3月預測值為19（千台）

月	銷售量（單位為千台）
2月	19
3月	18
4月	15
5月	20
6月	18
7月	22
8月	20





# 其他預測方法（85-95頁）

- 焦點預測法
- 散佈模型
- 趨勢分析技術
  - 線性趨勢方程式
  - 趨勢調整指數平滑法
- 季節性分析技術
  - 加法模型（趨勢＋季節）
  - 乘法模型（趨勢×季節）





# 關聯性預測技術

- 關聯性技術所使用的變數必須與感興趣之變數預測值相關。
- 關聯性技術
  - 重點在於建立出歸納預測變數(predictor variable)效果的方程式。
  - 主要的分析方法為迴歸(regression)。

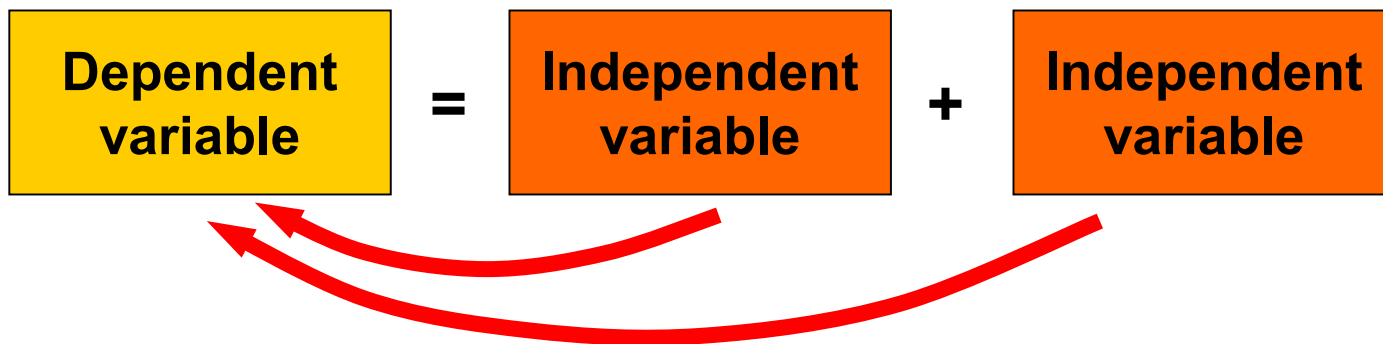


# 關於迴歸，你知道甚麼？



# 迴歸的概念

- *Dependent variable (response variable)* : the variable to be predicted
- It depends on *independent variables*



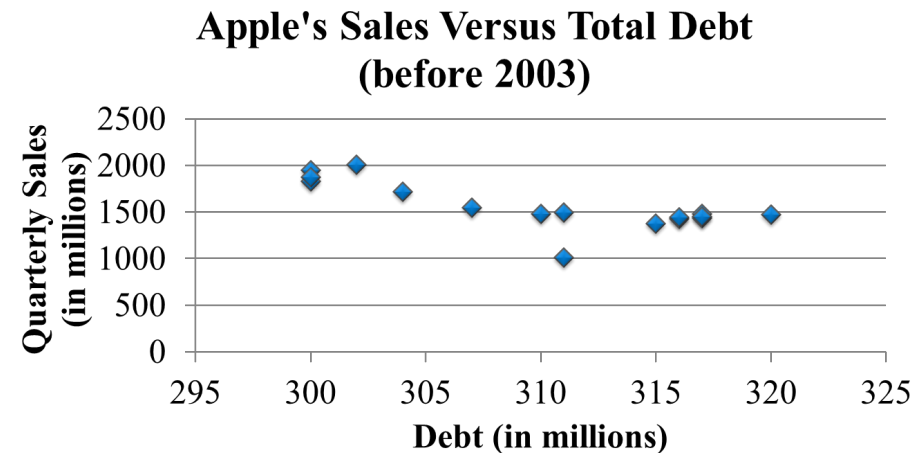
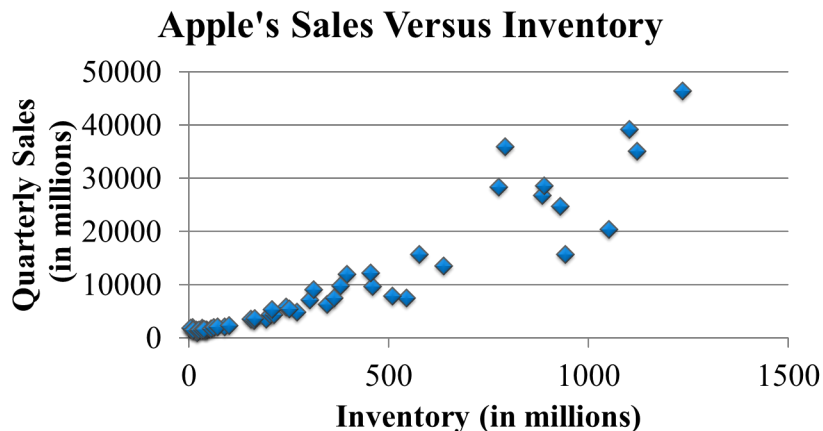


# 怎麼找獨立變數？



# 怎麼找獨立變數？

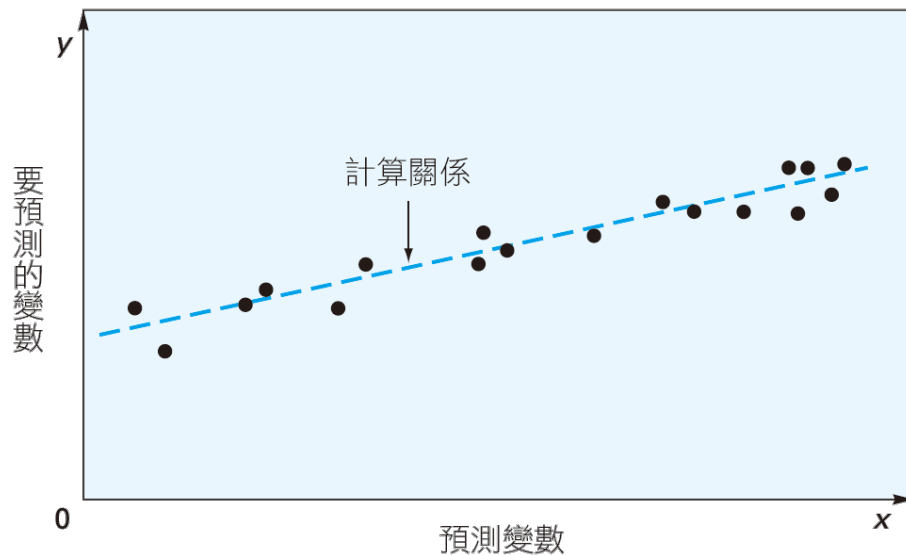
- Scatter Diagram (Plot)
  - The independent variable: on the  $X$  axis
  - The dependent variable: on the  $Y$  axis





# 簡單線性迴歸

- 最簡單與最廣泛使用的迴歸形式為兩個變數之間的線性關係。
- 線性迴歸的目的是求出一條直線方程式，使每個資料點與此線的垂直距離平方和最小。



注意：通常以y軸表示要預測的（相依）變數值，以x軸表示預測（獨立）變數值。



# 最小平方直線(least squares line)

$$y_c = a + bx$$

其中  $y_c$  = 預測（相依）變數  
 $x$  = 預測（獨立）變數  
 $b$  = 直線的斜率  
 $a$  為  $x = 0$  時， $y_c$  的值（即  $y$  截距的高度）

- 以下的方程式可以計算出係數  $a$  與  $b$ ：

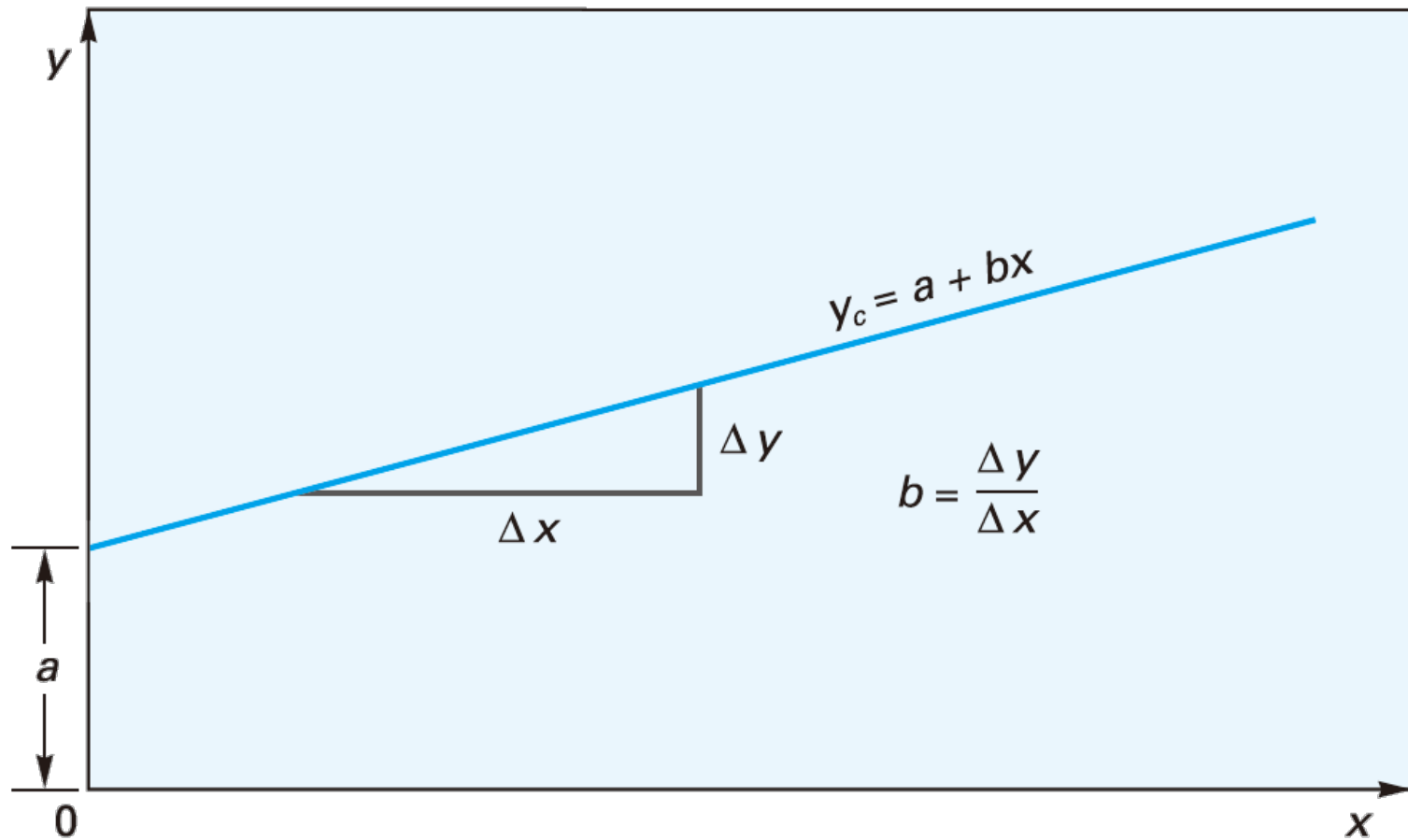
$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n} \quad \text{或} \quad \bar{y} - b\bar{x}$$

其中  $n$  = 成對觀察點的數目



# 最小平方法



直線的截距為  $a$ ，斜率為  $b$ 。





## 練習

- Develop a simple regression model that predicts the following firm's sales (in millions) based on its assets (in millions).
- Predict the sales of a quarter when the firm's asset value is 83 millions.

quarter		1	2	3	4	5	6	7	8	9
Assets		98	77	83	80	96	61	66	95	69
Sales		93	78	84	73	84	64	64	95	76



# 線性迴歸線之精確度

- 散佈在迴歸線附近的資料點之數量為預測線性迴歸線之精確度的一項指標。
- 散佈程度可以用**標準估計誤差 (standard error of estimate)**解釋。

$$S_e = \sqrt{\frac{\Sigma(y - y_c)^2}{n - 2}}$$

其中  $S_e$  = 標準估計誤差  
 $y$  = 每個資料點的 $y$ 值  
 $n$  = 資料點總數



# 決定系數

## (Coefficient of Determination, $r^2$ )

- Sum of the squares total  $SST = \sum (Y - \bar{Y})^2$
- Sum of the squared error  $SSE = \sum e^2 = \sum (Y - \hat{Y})^2$
- Sum of squares due to regression  $SSR = \sum (\hat{Y} - \bar{Y})^2$
- An important relationship  $SST = SSR + SSE$
- Coefficient of determination  $r^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$



## 練習

- Develop a simple regression model that predicts the following firm's sales (in millions) based on its assets (in millions).
- Find  $r^2$ ?

quarter		1	2	3	4	5	6	7	8	9
Assets		98	77	83	80	96	61	66	95	69
Sales		93	78	84	73	84	64	64	95	76



# 指標有效之條件

- 迴歸在預測的應用與指標的使用有關。
- 使指標有效的三個條件如下：
  1. 指標數據變動與變數變動之間的關係應該要有符合邏輯的解釋。
  2. 指標數據變動必須領先相依變數變動足夠的時間，使預測不至於過時。
  3. 兩變數之間應該有很高的相關性。



# 相關性(correlation)

- 衡量兩變數之間關係的強度與方向。
- 相關係數 $r$ 的範圍為 $-1.00 \sim +1.00$ 。
  - 相關係數為0即表示兩變數間幾乎沒有線性關係。

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

- 連結回scatter diagram



# 相關性(correlation)

- 相關係數的平方( $r^2$ )可用來衡量獨立變數對 $y$ 值變動所占的百分比。 $r^2$  值為0~1.00。
  - 愈接近1.00，所占的百分比愈高。
  - 若 $r^2$ 值相當高（例如，0.80 或以上），表示獨立變數是相依變數的優良預測值。
  - 若 $r^2$ 值為0.25或以下，表示為不良預測值。 $r^2$ 值介於0.25~0.80 之間，則為中等的預測值。



# 應用線性迴歸分析的要點

- 簡單迴歸分析的應用應滿足下列假設：
  - 在直線附近的變異是隨機的。
  - 在直線附近的偏差應為常態分配。
  - 只在觀測值的範圍內進行預測。
- 滿足上列假設後，為了得到最佳結果：
  - 經常將資料繪成圖形，驗證線性關係是否恰當。
  - 資料也許會受時間影響，檢查並繪出相依變數相對於時間的圖；若顯示出模式，則使用時間序列分析替代迴歸分析，或把時間當做多元迴歸分析(multiple regression analysis)的獨立變數。
  - 低度相關可能暗示有其他更為重要的變數。





# 應用線性迴歸分析的要點

- 迴歸分析的缺點包括：
  1. 簡單線性迴歸只能用在包含一項獨立變數的線性關係。
  2. 建立這種關係需要大量的資料－至少超過20 個觀測資料。
  3. 所有觀測值之權重皆相等。



# 非線性與多元迴歸分析

- 適用於包含一個以上的預測變數或不適宜的線性模型，而無法採用簡單線性迴歸時。
  - 存在非線性關係時，使用非線性迴歸。
  - 含一個以上的預測變數時，使用多元迴歸分析。

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

- Excel can help.



# 回到預測的精確度

- 預測**誤差(error)**是針對已知的期數，其實際值與預測值的差。

誤差＝實際值－預測值

$$e_t = A_t - F_t$$

- 預測誤差會以兩種稍微不同的方式影響決策。
  - 一種是在許多預測替代方案中選擇其一。
  - 另一種是評估預測技術的成功或失敗。



# 彙整預測精準度

- 精確度主要建立在預測的歷史誤差績效之基礎上。  
常用的衡量方法主要有三種：

- 平均絕對偏差 (mean absolute deviation, MAD)

$$\text{MAD} = \frac{\sum | \text{實際值}_t - \text{預測值}_t |}{n}$$

- 均方誤差 (mean squared error, MSE)

$$\text{MSE} = \frac{\sum ( \text{實際值}_t - \text{預測值}_t )^2}{n - 1}$$

- 平均絕對百分比誤差 (mean absolute percent error, MAPE)

$$\text{MAPE} = \frac{\sum \frac{| \text{實際值}_t - \text{預測值}_t |}{\text{實際值}_t} \times 100}{n}$$



# 練習：求解MAD、MSE 與MAPE

期數	實際值	預測值	(A - F)			
			誤差	誤差	誤差平方	(  誤差   ÷ 實際值) × 100
1	217	215				
2	213	216				
3	216	215				
4	210	214				
5	213	211				
6	219	214				
7	216	217				
8	212	216				



# 彙整預測精準度

- 總括來說，作業管理者必須決定歷史績效對反應能力的相對重要性，以及是否要使用MAD、MSE或MAPE來衡量歷史績效。
  - MAD最容易計算，但對所有誤差的權重相等。
  - MSE權重是根據其平方值，但也有較大誤差，所以會產生較多問題。
  - 當需要以相對的觀點來衡量誤差時，則使用MAPE。

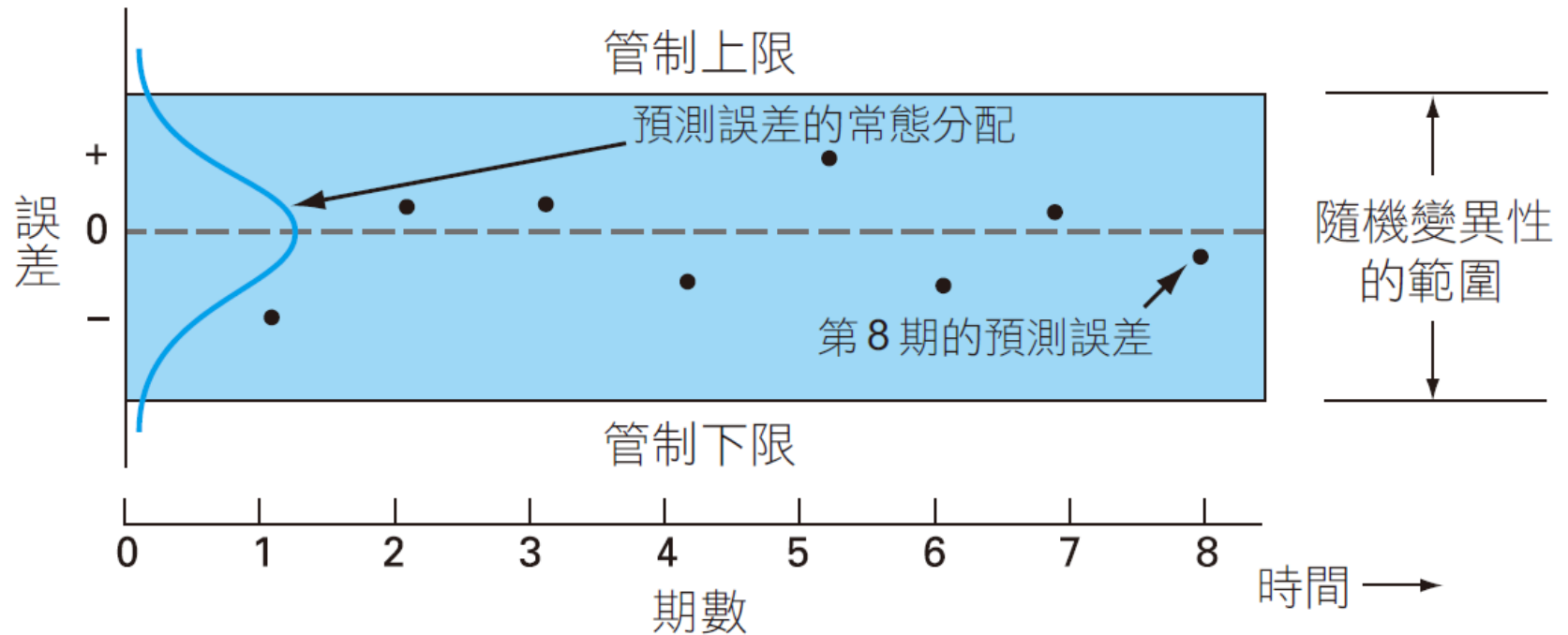


# 預測的檢視

- 預測誤差產生的原因：
- 偵測非隨機誤差的工具：
  1. 管制圖(control chart)。
  2. 追蹤訊號(tracking signal)。



# 管制圖(control chart)







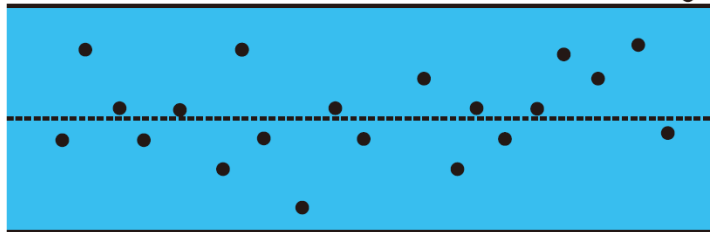
# 管制圖(control chart)

## ● 非隨機性範例

點落在管制界限外

管制上限

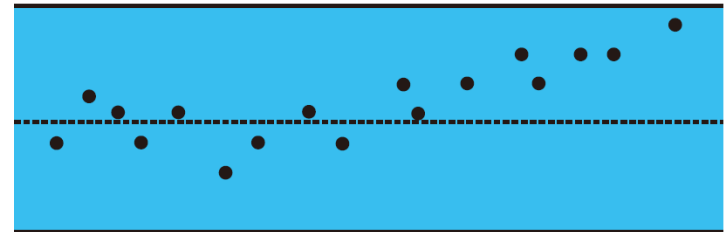
落於管制上限之上的誤差點



管制下限

趨勢

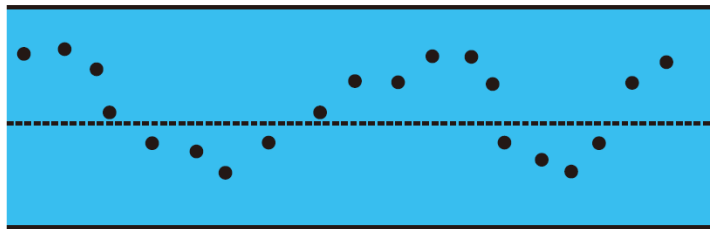
管制上限



管制下限

循環

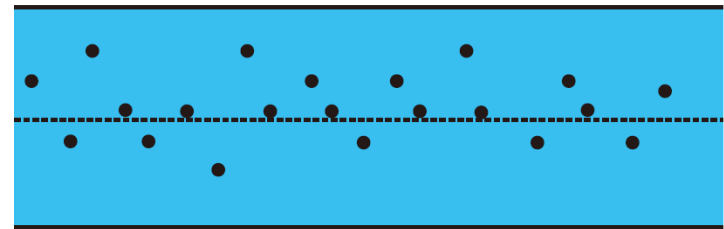
管制上限



管制下限

偏差（多數點位於管制界限的同一邊）

管制上限



管制下限



# 繪製管制圖

- 應計算MSE。誤差分配的標準差估計值即為MSE 的平方根。

$$s = \sqrt{\text{MSE}}$$

- 管制圖基本假設：
  - 當誤差為隨機時，誤差就會是常態分配，且平均值在0的附近。
  - 常態分配是約95.5%的值（即誤差）預期會落在  $0 \pm 2s$ （即  $0 \pm 2$  標準差）的界限內，約99.7%的值會落在  $0 \pm 3s$  的界限內。



# 繪製管制圖

- 求得管制上限與管制下限：

$$\text{管制上限} : 0 + z\sqrt{\text{MSE}}$$

$$\text{管制下限} : 0 - z\sqrt{\text{MSE}}$$

其中  $z$  = 來自於平均值的標準差

結合上述兩個公式，可得：

$$\text{管制界限} : 0 \pm z\sqrt{\text{MSE}}$$



# 選擇適合的預測方法之導引

## ● 選擇適合的預測方法

預測方法	過往資料的數量	資料模式	預測幅度	準備時間	人員背景
移動平均法	2~30 個觀測值	資料應該是穩定的	短期	短	少許經驗
簡單指數平滑法	5~10 個觀測值	資料應該是穩定的	短期	短	少許經驗
趨勢調整指數平滑法	10~15 個觀測值	趨勢	短期至中期	短	中等經驗
趨勢模型	10~20 個觀測值，每季至少有 5 個是季節性的	趨勢	短期至中期	短	中等經驗
季節性	足以看到 2 個高峰和離峰	處理循環與季節性模式	短期至中期	短至中等	少許經驗
因果迴歸模型	每個獨立變數有 10 個觀測值	能處理複雜模式	短、中、長期	發展時間長，完成時間短	豐富經驗



# 以預測的時間範圍分類預測因素

## ● 以預測的時間範圍分類預測因素

因素	短期	中期	長期
1. 頻率	經常	偶爾	不常
2. 總合程度	項目	產品家族	總產出 產品或服務種類
3. 模型種類	平滑 投射 迴歸	投射 季節性 迴歸	管理判斷
4. 管理者參與程度	低	中	高
5. 單次預測成本	低	中	高