販賣豫測과 統計的 方法의 利用

康 景 植*

1 序 論

1·1 豫測의 重要性

販賣豫測이란 將來의 一定時點에 있어서의 販賣可能量을 推定하는 것으로 이에 의하여 生産計劃에 직접적으로 영향을 주는 需要의 變動, 새로운 製品에 대한 消費者의 反應 등을 예상할 수 있게 된다. 明確한 豫測의 評價는 종속적인 決定을 考慮할 때 하는 것이 손쉬운 것이라고 할 수 있다.

따라서 販賣豫測과 生産量決定의 關係를 chart 로 나타내면 그림 1과 같다.

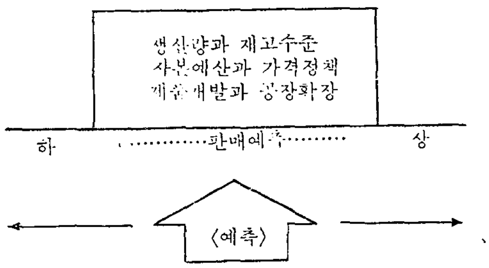


그림 1 판매예측과 생산결정

1·2 販賣豫測의 問題點

豫測에 있어서 그 優劣을 가름하는 尺度는 그 豫測의 正確性에 있다고 하겠는데 이 正確性은 原因과 結果間의 關係가 化學製品에 대한 消費者들의 反應이나 消費者들이 바라는 設計의 重點등과 같이 直接的인 境遇에는 比較的 쉽게 期待할 수 있다.

따라서 景氣變動, 物價變動, 政府의 經濟計劃이 製品販賣에 미치는 影響 등 經濟的인 領域에 이르르면 正確한 豫測은 相對的으로 어렵게 된다.

어떠한 複雜性을 가지는 變動要所들이 있는데 여기서 信賴할 수 있는 回答을 豫想한다는 自體가 벌써 非現實的이라는 것을 알 수가 있다.

그런데 아직도 一般 事業家들은 이러한 事實을 理解하지 못한 結果로 豫測에 對한 正確한 回答을 期待하려 들지 않고 있으며 또한 다른 豫測者도 그것을 試圖하려는 어리석은 짓을 하는 例를 흔히 볼 수 있다. 따라서 豫測을 한다는 것은 그들의 努力이 將來의 狀況判斷에 있어서 結論이라기 보다는 오히려 指示라는 것을 念頭에 두는 것만이 적어도 生産活動에서 일어나는 마찰과 病廢를 減少시킬 수 있을 것이고 豫測의 本來의 性質임을 認識할 수 있을 것이다.

1·3 Data 豫測資料의 出處

많은 企業들이 將來의 活動을 豫測하는데 있어서 豫測專門家들을 雇傭한다는 것은 容易한 일이 아니므로 비록 生産計劃이 한 사람의 管轄下에 있을지라도 適切한 豫測을 하기 위해 統計的인 分析을 할 수 있어야 하는데 그의 作業遂行에 있어 必要로 하는 모든 data를 使用하면 豫測하는데 많은 도움을 준다. 一般的인 data 즉 間接的인 資料로서 物價指數, 消費者의 性向등을 雜誌, 新聞 또는 國家의 公共機關과 貿易協會에서 發行되는 出版物에 依하여 參考로 할 수 있으며 이외에 直接的인 豫測의 data로서 消費者의 意見, 市場調査, 過去의 實績資料 등의 豫測을 爲한 資料로 삼을 수 있다.

1·4 時系列 分析

時間의 經過와 더불어 變動하는 統計單位를 時間의 順序에 따라서 配列한 것을 時系列이라고 한다.

時系列은 系統的 部分과 偶然的 部分으로 나누어진다.

系統的 部分은 長期的 變動傾向을 意味하는 趨勢와 廣意의 循環的 變動으로 나누어지고 廣意의 循環的 變動은 다시 正確하게 一定한 週期를 갖는 週期的 變動과 그렇지 않은 狹意의 循環的 變動으로 區分한다.

週期的 變動 가운데서 特히 一年을 週期로 하는것을 季節變動이라 부른다. 그리고 經濟的 時系列의 境遇에 있어서 狹意의 循環的 變動은 主로 經濟의 好氣景와 不景氣를 反應하는 것으로 생각하여 普通景氣循環 혹은 景氣變動이라 부른다.

時系列의 偶然的 部分은 不規則變動이라하며 그것

*明知大學校 工業經營學科 專任講師(1978年 10月 27日)

은 地震, 水害, 戰爭 등 突發的 變化의 影響을 받는 變動이다.

一般的으로 時系列에는 위와 같이 여러가지 相異한 變動이 同時에 包含되어 있으므로 주어진 時系列을 分析加工해서 各各 變動의 變動系列을 分別하는 것이 必要하게 된다.

이것을 時系列 分析의 中心問題로 趨勢變動과 季節變動을 測定하는 것이다. 趨勢變動과 季節變動을 구하면 그것 自體도 測定할 수 있을 뿐만 아니라 그것들을 除去하면 다른 循環變動이나 不規則變動을 求하는 데도 便利하기 때문이다.

2 時系列 計算과 方法

위에서 言及했듯이 時系列에는 趨勢變動(trend movement), 循環變動(cyclic movement), 季節變動(seasonal movement), 不規則變動(irregular movement)가 있는데 이들 中에서 實績資料에 依하여 趨勢變動과 季節變動에 關하여 統計的 技法을 使用하면 된다. 趨勢變動에서는 最小自乘法(least square method)과 指數平滑法(Exponential Smoothing)이 使用되고 季節變動指數(Seasonal Index)를 求할 시는 單純平均法(Simple Average method)과 平均法을 使用하면 된다.

여기서는 豫測의 手段으로 많이 使用되는 最小自乘法과 單純平均法, 移動平均法, 그리고 指數平滑法을 實例를 들어 考察하기로 한다.

例 1:

다음의 某會社는 지난 5年間 hand-operated(장도리), nail driving(못징)을 生産해 왔는데 工場은 지난 2年동안 工場의 生産能力(Production Capacity)에 거의 맞먹는 程度의 作業을 해왔다.

따라서 이 工場은 앞으로 生産을 爲한 豫測과 生産施設의 擴張計劃을 爲한 推定이 必要했다.

지난 5年間의 販賣實績은 表 1과 같다. 그리고 그림 2는 年間 販賣高의 "패턴"을 나타낸다.

表 1 分期別 販賣高(單位: ₩(10,000)

年	1964	1965	1966	1967	1968
分期 1	190	280	270	300	320
分期 2	370	420	360	430	440
分期 3	300	310	280	290	320
分期 4	200	180	190	200	200
合計	1,080	1,190	1,100	1,220	1,300

最小自乘法(least squares method)이 直線方向으로 흐르는 傾向이 있을때 普通 最適直線을 決定하기 爲하여 最小自乘法을 使用하는데 여기서 求해지는 直線은 모든 data point에 가까운 直線의 方程式이 된다

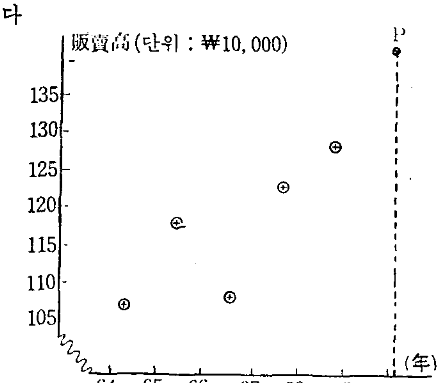


그림 2 연간판매고의 패턴

$\Sigma(Y-Y_F)^2$=最小(但 Y=實際値, Y_F=最小自乘의 값) 또한 훨씬 適切한 直線은 똑 같은 垂直距離의 合이 0일境遇 $\Sigma(Y-Y_F)=0$의 式이 얻어진다.

그러므로 直線은 方程式

$Y=a+bX$로 定義된다.

왜냐 하면 時系列分析 Y는 原點(base point)으로부터 每年增加分을 測定하는 時間에 있어서의 豫測된 價値이기 때문이다. 우리의 目的은 原點에서 X의 價値, 即 a와 直線의 기울기 b를 決定하는데 있다.

a, b를 決定하기 爲하여 두 가지 方程式을 使用하여야 하는데 첫째 方程式은 直線은 方程式에다 期間의 合과 a의 系數를 곱하므로써 얻을 수 있다.

$\Sigma Y=Na+b\Sigma X$

둘째 方程式도 마찬가지로 도출해 낼 수 있다.

b의 系數 X에 各 期間을 곱한다음 이를 모두 합하면 다음의 方程式을 얻는다.

$\Sigma\times Y=a\Sigma X+b\Sigma X^2$

이와 같이 얻어진 두 方程式을 正規方程式(Normal Equation)이라고 한다.

이 式을 連單立 하여 풀면 ΣY, ΣX, ΣXY, ΣX^2을 代入하여 a와 b를 計算할 수 있다. 따라서 最小自乘法의 豫測을 爲한 直線의 方程式은

$Y_F=a+bX$이다.

例題; 趨勢를 爲한 最小自乘法 適用의 事例 表 1의 data를 使用하여 直線의 方程式을 求해 보자.

最小自乘法의 相異한 기법을 보이기 위해서 原點으로서 1964年을 使用한뒤 1966年을 使用하는 法을 보이겠다. 아래와 같이 표를 만들면

Year	Y	X	X^2	XY
1964	108	0	0	0→原點
1965	119	1	1	119

1966	110	2	4	220
1967	122	3	9	316
1968	130	4	16	520
合計	589	10	30	1,225

을 얻을 수 있다.

이것을 正規方程式에 代入하면

$589 = 5a + 10b$

$1,225 = 10a + 30b$

이 方程式을 풀면 $a=108.4$, $b=4.7$을 얻을 수 있다.

1966年을 原點으로 使用하면 다음 表를 만들 수 있다.

Year	Y	X	X^2	XY
1964	108	−2	4	−216
1965	119	−1	1	−119
1966	110	0	0	0→原點
1967	122	1	1	122
1968	130	2	4	260
合計	589	0	10	47

따라서

$a = \Sigma Y/N = 589/5 = 117.8$

$b = \Sigma XY/\Sigma X^2 = 47/10 = 4.0$

을 구할 수 있다.

豫測方程式은 直線의 方程式에다 a, b의 값을 代入하여 求할 수 있다.

1964年을 原點으로하여 5年後인 1969年의 豫測方程式은

$Y_F = 108.4 + 4.7X$

따라서 $F_{1969} = 108.4 + (4.7 \times 5)$

똑같이 1966年을 原點으로 하였을 境遇

$Y_F = 117.8 + 4.7X$

의 豫測方程式이 얻어지는데 1969年의 豫測은 $F_{1969} = 117.8 + (4.7 \times 3) = 131.9$

그러므로 위에서 보아온 두 豫測方程式의 計算을 比較해 보면 基本的인 問題는 原點이 다르다는데 있다 그러나 豫測値는 같다는 것을 알 수 있다.

單純平均法(Simple Average Method)

直線 $Y=a+bX$에서 b가 0일때 이 直線은 변수 X에 平行이 된다.

다음 期間동안 豫測은 모든 Y値의 單純平均에 衣해서 行하여진다.

따라서 $Y_F = \Sigma Y/N \cdots\cdots (1)$

趨勢豫測을 위한 單純平均의 計算은 最小自乘의 特別한 境遇에 適用되는 技法이다.

平均値計算은 全般的인 趨勢下에 일어나는 週期的 變動에 關聯을 가지고 있다. 定義에 依하면 週期的 變動은 單純히 1年동안의 變動에 限하고 있으며 週期的 變動의 "패턴"을 決定하기 위해 1年間의 여러 期間 동안의 資料를 蒐集해야 한다.

月間 혹은 分期의 記錄이 가장 一般的으로 使用된다.

年間 實際的인 數字를 나타내는 資料가 有効할때 1年間의 各 期間의 算術平均은 數年間의 循環的 變動의 効果를 平均시키는 傾向이 있다.

分期間, 月間에 依한 資料를 配列한 뒤 各 期間의 Y의 代表値를 合하고 이것을 年數 N으로 나눈다.

各 期間동안의 最終平均은 增加趨勢 効果가 言及되지 않을 境遇 다음 期間동안의 豫測이 된다.

趨勢가 重要할때 單純平均은 全般的인 成長 또는 低下를 수정할 수 있다.

週期的 變動의 "패턴"은 趨勢가 增加하거나 減小한지라도 比較的 一定하다.

變動 "패턴"은 各 期間의 單純平均을 平均의 合으로 나눔으로써 쉽게 求할 수 있다.

結果數値는 各 期間동안 豫測되는 活動量의 百分率이 된다.

이러한 週期的 變動指數는 各 期間의 百分率에다 每年 趨勢豫測을 곱함으로써 販賣나 需要單位를 變換시킨다.

그러면 이것을 例를 들어 考察하여 보자.

예제 ; 週期的 變動指數에 對한 單純平均 法適用의 事例 앞에서 言及한 定義에 依하면 表 1에 依하여 다음 表를 作成할 수 있다.

Year	Q_1	Q_2	Q_3	Q_4	Annaol
1964	190	370	300	220	1,080
1965	280	420	310	180	1,190
1966	270	360	280	190	1,100
1667	300	430	290	200	1,220
1968	320	440	320	220	1,300
合計	1,360	2,020	1,500	1,010	5,890
平均	272	404	300	202	2,945

이것을 圖表로 그리면 그림 3과 같다.

各 期分의 平均은 週期的 變動指數로 變換되어야 한다.

따라서 各期分의 指數는 全體 平均에 依하여 各期分의 平均을 나눔으로써 얻어지는 常數이다.

$I_{Q1} = 272/294.5 = 0.92$

$I_{Q_2}=404/294.5=1.37$

$I_{Q_3}=300/294.5=1.02$

$I_{Q_4}=202/294.5=0.69$

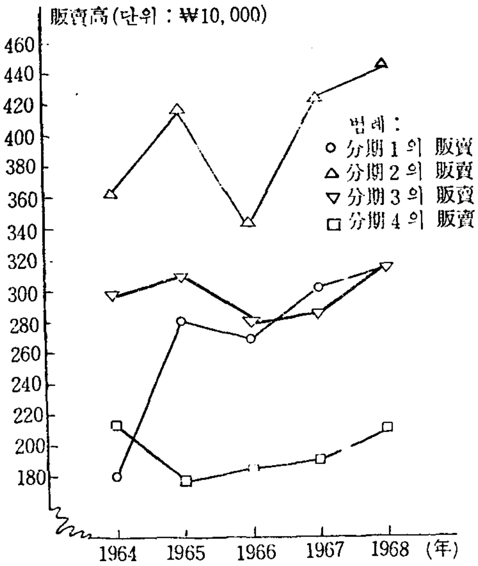


그림 3 분기별 판매추세

이와 같이 計算된 指數는 明年에 對한 各 期分의 販賣量을 推定하는데 使用될 수 있다.

最小自乘法에 依해 求해진 1969年의 豫測値를 適用시킴으로써 다음과 같은 各 期分의 販賣量을 算出할 수 있다.

$F_{Q_1}=(1,319,000/4)\times 0.92=303,000$

$F_{Q_2}=(1,319,000/4)\times 1.37=452,000$

$F_{Q_3}=(1,319,000/4)\times 0.12=336,000$

$F_{Q_4}=(1,319,000/4)\times 0.69=228,000$

總計 1,319,000 : (1969年의 豫測値

移動平均法(Moving Average Method)

移動平均法에 依한 豫測은 지난 期間의 數字를 通하여 data point의 合에 依하여 얻어지는데 이 數字는 週期的 變動을 다루기 爲하여 1年을 單位로 한다.

또한 移動平均法은 繼續的인 計算을 要함으로써 單純平均과는 區別된다.

例를들면 12個月 移動平均法이 1968年 1月부터 12月까지의 平均需要를 나타낸다고 할때 次期 移動平均은 1968年 2月부터 1969年 1月까지의 需要를 包含한다.

前者의 境遇에 있어서 移動平均은 6月 30日이나 7月 1日의 需要를 나타내며, 後者의 境遇에 있어서는 7月 30日이나 8月 1日의 需要를 다타낸다. 두 價値의 平均은 7月의 需要를 다타낸다.

가장 最近에 觀測된 data의 數에 依한 移動平均이라 할지라도 만약 data pattern이 比較的 一定치 않다면 次期에 있어서의 좋은 豫測이 될 수 없다.

移動平均을 參考로한 週期的 變動指數는 豫測을 容易하게 한다.

指數値는 실지의 需要를 그 期間동안 中心이 되는 移動平均으로 나눔으로써 計算할 수 있는데 좀더 信賴할 수 있는 指數는 普通 그 期間동안 여러개의 指數値를 平均함으로 얻을 수 있다.

따라서 豫測은 한期間동안 中心이되는 移動平均과 그 期間동안의 指數値의 곱이라 할 수 있다.

그러면 例題를 들어 說明하기로 한다.

例題 ; 週期的 變動指數에 對한 移動平均法 適用의 例

表 1을 利用하여 1年을 4項移動平均法에 依하여 中心移動平均(Centered Moving Average)과 週期的 變動指數를 求하면 表 2와 같고 週期的 變動指數의 조정計算은 表 3의 計算으로 다타냈다.

<表 2>

年	分期	販賣高 (₩ 10,000	4分期間 移動平均	移動平均	指數 週期
1964	1	190			
	2	370			
			270		
	3	300		281	1.07
			292		
	4	220			
1965	1	280	305	298	0.74
	2	420	307	306	0.91
	3	310	297	302	1.39
	4	180	295	296	1.04
1966	1	270	280	287	0.63
	2	360	273	276	0.93
	3	280	275	274	1.32
	4	190	283	279	1.00
1967	1	300	300	286	0.66
	2	430	303	301	1.00
	3	290	305	304	1.42
	4	200	310	307	0.92
1968	1	320	312	311	0.64
	2	440	320	316	1.01
	3	320	325	322	1.37
	4	220			

그런데 週期的 變動平均指數를 適用하기에 앞서

〈表 3〉 調整된 週期的 변동지수의 計算

	Q_1	Q_2	Q_3	Q_4
			1.07	0.74
	0.91	1.39	1.04	0.63
	0.98	1.32	1.00	0.66
	1.00	1.42	0.94	0.64
	1.01	1.37		
合計	3.90	5.50	4.05	2.67
平均 週期 指數	0.975	1.375	1.0125	0.6675
調整된 週期 指數	0.97	1.37	1.00	0.66

檢討해야 할 두가지 點이 있는데 첫째, 期間指數의 平均이 1.0이 되어야 한다.

表 5에서 보면

(0.9750+1.3750+1.0125+0.6675)1/4=4.03/4=1.0075

그러므로 指數는 調整되어야 하며 그렇지 않으면 期分豫測의 合이 0.75%程度에서 每年 豫測을 超過하게 된다.

둘째, 分期別 指數에 있어서 明確한 趨勢에 注意하여야만 한다.

表 3에서 Q_1은 增加하고 있으나 Q_3은 오히려 減小하는 趨勢를 나타내고 있다. 최종단계로서 豫測을 해볼 때 중심 移動平均에다 各各의 週期的 變動指數를 곱하면 次期의 各 期分間의 豫測値를 얻을수 있다.

Q_1 1969=316×0.97=307 또는 307,000

Q_2 1969=322×1.37=441 또는 441,000

指數平滑法(Eponential Smoothing Method).

어떠한 量的인 豫測方法일지라도 需要「패턴」에 있어서의 變動을 豫測함으로써 원활하게 역할을 한다는 것을 앞에서 언급한 여러가지의 技法에 依하여 알아 왔는데 특히 指數平滑法에 있어서는 因子 α를 副加함으로써 傾向線의 特性을 統制管理할 수 있다.

指數法은 어떠한 時系列 豫測技法으로 適用됨에 있어서는 특히 平均値의 計算方法에도 적용될 수 있는데 어떠한 경우에도 관련시켜서 檢討해야 한다. 따라서

$$F_n=\alpha Y_{n-1}+(1-\alpha)F_{n-1}\cdots\cdots(2)$$

의 方程式을 얻을 수 있다.

이式을 변경하면,

$$F_nF_{n-1}+\alpha(Y_{n-1}-F_{n-1})\cdots\cdots(3)$$ 이 된다.

단, F_n=차기의 예측치

F_{n-1}=전기간의 예측치

α=평활(Smoothing 常數($0\leq\alpha\leq1$)

Y_{n-1}=전기간의 實際價値

指數평활法에 의한 豫測은 公式에서 보여주는 바와 같이 前期間동안 豫測値의 實際値와 差에다 因子 α를 곱해준 다음 前期間의 豫測値를 더해준 것이다

따라서 이것을 풀려면 우선 새로운 豫測을 하기 前에 α의 價値와 前期의 豫測을 決定해야만 한다.

實績資料가 있을 경우엔 初期値 F_{n-1}는 單純平均에 依해서 求할 수 있으며 또한 移動平均法에 依해서도 求할 수 있다.

그밖의 경우엔 F_{n-1}은 推測에 依해서 얻어진다.

α値의 表擇을 위한 理論的原理는 移動平均에 있어서 期間數 N을 選擇할때 使用되는 原理와 유사하다.

一般的으로 α는 0.1과 0.3사이의 값이 된다. 따라서 α와 N와의 사이에는 $\alpha=2/(N+1)$인 관계가 있으므로 어떤 N値에 만족할때 α는 쉽게 計算할 수 있다.

그런데 N가 증가할때는 F_n의 式은 變形이 되는데, F_1, F_2, F_3, 등의 式을 보이면

期間 1 : $F_1=\alpha Y_0+(1-\alpha)F_0$

期間 2 : $F_2=\alpha Y_1+(1-\alpha)F_1$

$=\alpha Y_2+(1-\alpha)[\alpha Y_0+(1-\alpha)F_0]$

期間3 : $F_3=\alpha Y_2+(1-\alpha)2F_2$

$=\alpha Y_2+\alpha(1-\alpha)Y_1+(1-\alpha)^2[\alpha Y_0+(1-\alpha)F_0]$

따라서 一般式은

$$F_n=\alpha Y_{n-1}(-\alpha)F_{n-1}$$
$$=\alpha Y_{n-1}+\alpha(1-\alpha)Y_{n-2}+\alpha(1-\alpha)^2Y_{n-3}+\cdots\cdots+(1-\alpha)^nY_{n-n}$$

단, F_n=n기간 동안의 豫測値

Y=實績値

Y_{n-n}=豫測最初値 즉

$Y_{n-n}=F_0$이다.

위의 함수는 平滑常數 α에 依한 過去의 모든 資料의 線形의 結合이다.

$\alpha=0$일경우 F_0이므로 豫測値는 存在하지 않으며 $\alpha=1$일경우 豫測値는 實際値에 가장 가깝게 된다.

그러면 指數平滑法에 依한 豫測을 例를들어 설명하기로 한다.

例題 : 豫測을 위한 指數平滑法 適用을 事例 表 1을 使用하여 F_n을 求하는 경우 첫째 $\alpha=0.2$라 할 때,

$F_{1969}=(0.2)130+(0.2)+(0.8)122+(0.2)(0.8)^2110+(0.2)(0.8)^3119+(0.8^4108$
$=(0.2)130+(0.16)122+(0.13)110+(0.10)119+(0.41)108=116$ 또는 1,160,000

둘째 $\alpha=0.8$ 이라 할때

$$F_{1969}=(0.8)130+(0.8)(0.2)122+(0.8)\times(0.2)^2 110+(0.8)(0.2)^3 119+(0.2)^4\times 108$$
$$=(0.8)130+(0.16)122+(0.03)110+(0.01)119+(0.0)\ 108$$
$$=128 \text{ 또 } 1,280,000$$

이러한 α값의 相異한 차이에 依한 豫測으로부터 α에 對한 신중한 選擇의 重要性이 있음을 알 수 있다.

초기 豫測値 F_n은 이외에 다른 平均法에 依하여 얻을 수 있다.

따라서 1967年의 Q_1 單純平均과 移動平均을 기ㅊ
로 하여 1969年의 1期分의 販賣高를 豫測할 수 있다

이때 $\alpha=0.15$로 잡아 安定度와 感度의 貸借를 앎
앤다고 가정하여 F_n을 求해 보자.

求하는 方法과 求한 값 F_n은 表 4와 같다.

表 4에 있어서 豫測値는 크게 다르다고 할 수 있ㄷ

이러한 현저한 차이는 초기 예측의 선택이4차후ㅇ
연속적인 豫測에 多分이 계속적으로 영향을 미친ㄷ
는 것을 말해주는 것으로 초기 豫測方法엔 상관없ㅇ
指數平滑法이 계속 여러기간동안 適用될 때엔 결ㄱ

<表 4> 單純平均과 移動平均을 基礎로한 分期別豫測에 對한 指數平滑法의 適用

豫測 方程式 : $F_n=F_{n-1}+0.15\ (Y_{n-1}-E_{n-1})$

單純平均	移動平均
$F_{n-1}=F_{Q1\cdot1967}\dfrac{190+280+270+300}{4}=260$	301(表 2에 依한 非調整移動平均 320
$Y_{n-1}=Y_{Q1\cdot1969}$ 320	301+0.15(320−301)
$F_n=F_{Q1\cdot1968}260+0.15(320-260)=269$	=304 혹은 $304,000
혹은 $269,000	

적으로 豫測의 값은 近似値에 가까와질것으로 생각되며 이 方法은 實際的으로 豫測者들을 고무시키는 方法이 될 것이다.

3 時系列豫測의 對照標準

지금까지의 豫測方法에 對해서 생각해 왔는데 이러한 種類의 豫測이 實際로 얼마만한 確實性을 갖고 있는지에 關해 直線性의 檢定과 反復的인 檢定方法에 依해서 檢定해 보기로 한다.

直線性의 檢定: 趨勢에 맞는 方程式의 檢定은 그 檢定에 對한 標準偏差 S_y에 依해서 檢定할 수 있다.

豫測된 點의 선형에 관한 실지적인 data point의 오차에 관한 측정은 다음과 같다.

$$S_y=\sqrt{\sum(Y-Y_F)^2/V}\cdots\cdots(4)$$

단 : Y=過去의 資料點

Y_F=豫測點

V=自由度

S_y를 使用하기전에 계산에 포함된 가정을 조사해보면 첫째, 豫測値 Y_F는 X시점에 있어서 어떤점과 관련이 있는 모든 Y관측의 평균이라고 가정했다

그러므로 X와 관련이 있는 Y치는 그림 4의 표와 같은 분포와 분산을 갖고 있다.

이러한 가정의 意味는 뒤에 나오는 상관 분석에서 명확히 알수 있는 것이지만 분산 σ^2의 평가는 S_y^2에 依하여 구할 수 있으며, 이것은 data가 많으면 많을수록 더 명확한 값을 기대할 수 있는것이다. 따라서 data가 30일때 정규적으로 Y치가 分布되어 있
다고 가정하자.

이러한 가정하에서 관측의 95%를 $\pm 2S_y$의 표준편ㅊ
사이에서 기대한다고 하면, data point가 30보다 적
을경우 student'st 分布를 使用하는 것이 더 좋을 ㄱ
이다.

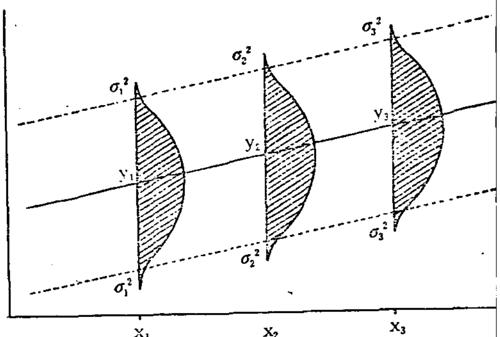


그림 4 직선상의 추정

이러한 가정은 관측중 95%의 신뢰도를 포함한다

이를테면 sloping line예측이 13번 관측을 기초ㅡ
한다면 자유로이 S_y를 계산하는데 使用될 것이다.

Studen'st 分布表로부터 95% band는 $Y_F\cdot S_y$ㄱ
된다.

표준편차의 평가는 과거의 data line에 의한 알ㅣ
는 line의 合理的測定을 하는데 있다. 그것만으로ㅡ

line이 將來에 얼마나 잘맞는지 신뢰할 수 있는 측정은 못된다. 따라서 판단력이 더 있어야 하겠다.

판단을 용이하게 하기 위한 것으로 도표상에 새로운 data를 나타낼 수 있다.

표준편차 관리한계에 관해서 나타난 pattern은 예측방정식의 持續價値에 대한 단서를 마련해 주고 있다.

그러면 實際的으로 어떻게 使用하는지를 實例를 들어 설명해보기로 한다.

例 2;

level-line에 대한 자유도가 4[5−1]이며, 反面에 sloping-line의 자유도가 3[5−2]이라고 생각하자.

여기서 level-line과 sloping-line方程式에서는 일

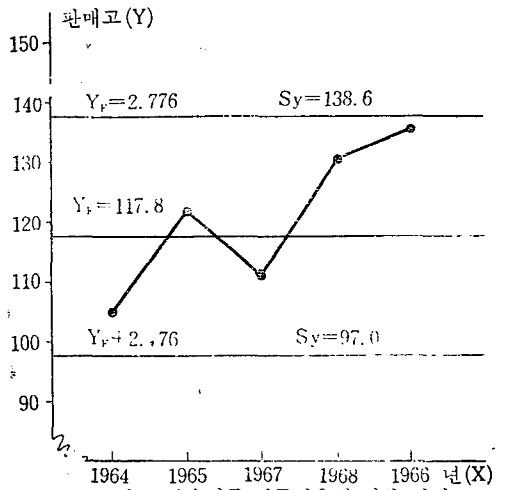


그림 5 단순이동 평균예측에 대한 관리도

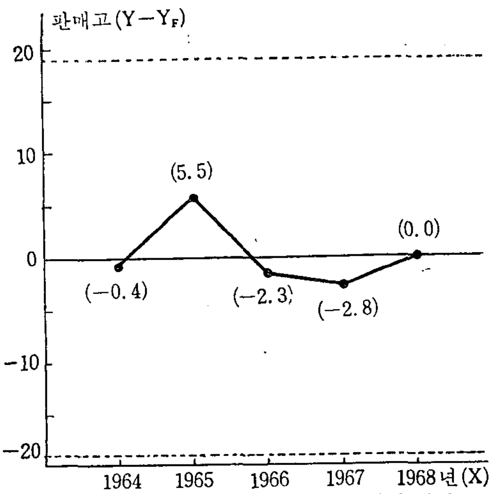


그림 6 최소자승 예측에 대한 누적적 관리도

어나지 않는다.

표준오차 계산을 기초로 한 통계관리 한계는 여러가지 상이한 관리도 公式으로 利用될 수 있다.

그림 5에서 실지관측은 사실대로 plot되었다.

適正한 직선이 level 또는 sloping이나 curve가 아니라면 관리한계는 slope 또는 curve에 평행하게 놓여진다.

그림 6에서 보는바와 같이 2차적인 說明은 예상치에 의한 실지가치의 누적편차($Y-Y_F$)를 따른다.

그림 1-6에서 보는바와 같이 2차적인 마지막 point는 예측방정식이 $\Sigma(Y-Y_F)=0$를 만들기 위해서 도표상에 보여준 모든 data의 使用에 依하여 개발된 예측방정식이기 때문에 0 line에 表示된다.

方程式이 모호하게 各 期間이 설명된다면 將來의 綜合偏差는 0 line의 한쪽면에 치우치게 될 것이다.

sloping-line豫測에 對한 lower standard error는 시각적으로 명확성을 확립하는 경향이 있다. 그림 5와 그림 6에 도시되어 있는 도표는 한계선을 기초로 하여 data의 95%를 포함하는 그림표이다.

data의 주어진 퍼센트를 설명하는 범위는 分布(모양) 형식의 함수임과 동시에 표준편차 함수이다.

例 2에서 표준편차 S_y의 계산을 했고 Student'st는 data point의 limited number에 依한 적절한 分布라는 것도 가정해 왔다.

예측방정식의 자유도가 서로 다르다는 것을 인식하고 Student'st표에 의한 各各의 범위는

$\pm 2.776S_y=\pm 20.8$(단순평균예측)

$\pm 3.182S_y=\pm 18.8$(최소자승법예측)

이 됨을 알수 있다.

또한 도표에서 도면에 표시된 점들은 특정지어진 data의 예측방정식에 관한한 한계안에 또는 관리선 아래 右에 임의로 표시될 것이다.

점들이 관리선의 한쪽에만 모여있거나 관리한계선의 밖에 있다면 예측에 경계가 있어야 한다는 신호이다.

따라서 예측자 즉 Industrial Engineer는 어떠한 信號라는 것과 Clue도 무시할 수 없으며 항상 도표상의 변화의 예에 주시하여야만 한다.

가끔 예측을 하기 위해서는 여러기간 동안의 누적된 수요를 요구하게 된다. 누적된 수요에 대한 표준편차의 대략측정은 예측된 수요가 누적되는 기간의 수 n에 依한 $S_y\sqrt{n}$에 依하여 計算된다.

앞에서 도출한 바 있는 최소자승법에 의한 예측에 적응함으로써 2年間 누적된 수요에 대한 표준편차

는 다음과 같이 계산된다.

期間		예측	누적예측	표준n
年	(n)	수요 ΣF_y	수요 ΣF_y	오차 $5.9\sqrt{n}$
1968	현재	127.2	—	5.9
1969	+1	131.9	131.9	5.9
1970	+2	136.6	268.5	8.2

이와같이 1970년에 대한 누적된 실지수요는 100번중 다섯번 정도는

$$268.5\pm3.182(8.2)=268.5\pm26.1$$
$$=294.6\sim242.4$$

이범위 밖에 우연히 표시될 것이라는 것을 알 수 있다.

누적된 수요에 대한 관리한계의 도표는 장기간에 걸친 예측에 대한 증가되는 발산을 명백히 보여줄 것이다.

일반적으로 관리도는 간단하면서도 유용한 실험 도구인 것이다.

관리도는 근본적으로 쉽게 지속되는 경고도구의 역할을 한다.

과거의 자료가 도면에 표시될 경우에 통계적인 안정상태의 지표를 나타내게 되는 것이다.

새로운 data point를 도면에 나타낸다면 現在의 活動이 과거의 모형을 따르고 있는지를 자세히 보여주는 것이 될것이다. 또한 어느정도까지는 긴급히 나타난 새로운 모형은 예측방정식을 어떻게 적절히 변경시킬 수 있는가를 제기한다.

도표의 중요한 양상외의 특이한 것은 近來의 實際値와 豫測値간의 불연속적인 比較에 있다.

따라서 data의 變化를 인식할 시기는 그것이 發生하고 있는 때에 있다.

反復的인 檢定 移動傾向을 豫測하는데 直接的으로 使用되는 것은 實地需要에 해당하는 價値를 나타내게 되는데 있다. 따라서 이러한 條件은 平均과 週期的 指表와 연관됨으로써 돋보이게 된다.

必然的으로 이러한 의지(vecourse)는 거의 現在의 data의 용도를 잃은 것이다.

즉 기간수의 半에 해당하는 중심 예측은 평균에 사용된다.

크기 N에 따라서 移動平均의 計算이 복잡하고 지루함을 나타냄을 알 수 있다.

平滑要因의 用途는 단지 최근 data 관측과 앞서의 예측을 利用함으로써 탁상사무와 計算을 단축하는데 도움이 될 것이다.

管理圖는 또한 移動平均과 함께 使用되며 管理限界는 다음과 같은 次例에 依하여 計算된 $M\bar{R}$, 즉 平均移動範圍(average moving range)를 基礎로 定해진다.

1. 移動平均에 依한 예측치와 실지 data사이의 차이를 얻어야 하며,
2. step 1에 依한 數(MR)의 차이의 절대적인 價値를 決定해야하며,
3. step 2에서 얻어진 價値의 平均을 내야한다.

$\pm2.66M\bar{R}$에 의한 限界는 우연히 0.3%보다 적은 이들 限界를 넘는 확률에 있고 3σ원칙을 기초로 해야한다.

그러한 管理圖는 過去의 data가 統計的으로 安定할지 어떤지 現 data가 과거의 모형을 따르고 있는지를 말하는데 도움이 될 것이다.

例 3;

moving range 관리도 작성을 위한 준비작업 단계, 有能한 豫測者는 훌륭한 豫測을 한 뒤에도 그의 豫測을 믿을 수는 없다. 그는 계속 統計的으로 어떤 意味있는 變化가 있는가를 찾아내기 위하여 實際的인 사진의 發生에 對해서 그의 끊임없는 豫測을 하여야만 한다.

移動平均을 使用한 會社에 대한 예측자는 表 5에 나타난 계산을 기초로 하여 管理圖를 發展시킬수 있으며 이 表의 맨 오른쪽의 열은 moving rang이다.

따라서 조정된 移動平均이 過去의 data에 얼마만큼 接近하고 있는지를 나타내고 있다.

또한 將來의 data가 얼마나 잘 맞을 것인지의 標準을 세워야 한다.

이러한 標準이 平均移動範圍인 것이다.

$$M\bar{R}=\Sigma MR/N=235/15=15.67 \text{ 또는 } 15.667$$

管理圖表에서 $M\bar{R}$은 앞에서 論義한 S_y에 相應한다.

S_y 즉 3σ에 一致하는 限界를 정하기 위하여 2.66을 平均移動範圍에 곱해야 한다.

그러므로 基本的인 data에 對한 移動平均圖表의 管理限界는 다음과 같다.

$$\pm2.66M\bar{R}=\pm2.66\times15.67=\pm41.68$$

이러한 한계는 平均 line 즉 管理線○의 양쪽에 位置해 있음을 나타내며 移動平均値는 그것들이 쓸모 없도록 圖表上에 plot해야할 것이다.

<表 5> 調整移動 平均販賣豫測에 依한 MR의 計算

年	分期	Y	中心移動平均	調整週期指數	Y_F	$Y-Y_F$	MR
1964	3	300	281	1.00	281	19	6

	4	220	298	0.66	196	24	40
1965	1	280	306	0.97	296	−16	21
	2	420	302	1.37	415	5	11
	3	310	296	1.00	296	16	26
	4	180	287	0.66	190	−10	11
1966	1	270	276	0.97	269	1	16
	2	360	274	1.37	375	−15	16
	3	280	279	1.00	279	1	0
	4	190	286	0.66	189	1	5
1967	1	300	301	0.97	294	6	9
	2	430	304	1.37	415	15	32
	3	290	307	1.00	307	−17	13
	4	200	311	0.66	206	−6	19
1968	1	320	316	0.97	307	13	10
	?	440	322	1.37	441	3	
						合計=235	

結果的인 plot pattern의 說明은 例題 8에서 言及했던 것과 같다.

4 相 關

相關分析은 變數사이의 關係程度를 調査하는 것을 말하는 것이다.

따라서 時系列計算에서 豫測된 標準誤差를 가지고 이것을 取扱하기 위하여 獨立變數 X年과 從屬변수 Y販賣量에 對해서 우리의 豫測이 얼마나 잘 어울리는지를 판단하는 S_y를 決定해 왔다.

이러한 事實을 出發點으로 해서 우리는 다른 有用한 통찰력을 얻기 위해서 더 깊게 상관관계를 탐구할 수 있다.

相關關係는 時系列分析에 局限되는 것이 아니며 變數에 關係하여 회귀선의 硏究에까지 適用될 수 있다

우리의 豫測方程式은 時間과 相關에 따라 變하는 販賣量은 또한 國民總生産, 가계부, 人口成長, 혹은 다른 經濟指表와도 연관될 수 있다.

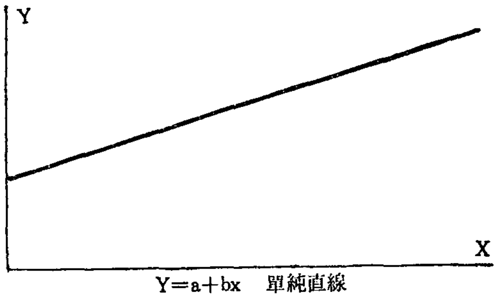
따라서 相關의 種類를 검토해 보면 첫째 單純相關은 2가지 變數사이의 상관을 表示하며 그림 7에 나타낸 회귀선과 관련이 있으며 둘째 다원상관은 2가지 變數以上 사이의 關係를 고찰할때 생기는 것으로 그림 8의 圖表는 두개의 獨立變數와 한 決定係數, 회귀직선에 있어서 時點의 分散은 세계의 제곱의 合으로 이루어지는 특징을 갖고있다.

$$\Sigma(Y-\bar{Y})^2=\Sigma(Y_F-\bar{Y})^2+\Sigma(Y-Y_F)^2 \quad (1\text{-}5)$$

단, $\Sigma(Y-\bar{Y})^2$은 全變動

$\Sigma(Y_F-\bar{Y})^2$은 用量間變動

$\Sigma(Y-\bar{Y})_F)^2$은 用量內變動(誤差)



Y=a+bx 單純直線

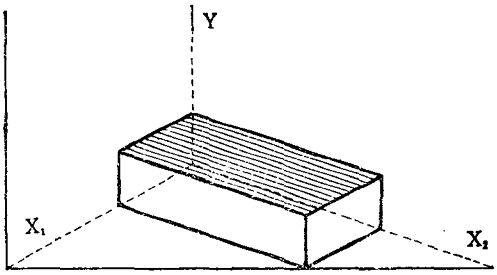
Y=a+bx+cx² 單純曲線

그림 7 1개 독립변수에 대한 회귀선

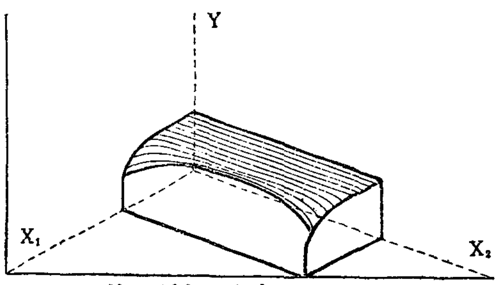
全變動 ; 그림 9에서 보여주는 바와 같이 편차는 평균($\bar{Y}=\Sigma Y/N$)과 동점과의 수직거리를 나타낸다. 全變動은 從屬變數의 全變動의 測定이며 두가지 部分으로 나누어진다.

즉 用量間變動과 用量內變動으로 나누어진다.

用量間變動 ; 회귀직선에 依해서 說明된 變數는($Y_F-\bar{Y}$)로 나타내어 지는데 그림 12에서 나타나는 바



Y=a+bx₁+cx₂ 多元直線



Y=a+blogx₁+clogx₂ 多元曲線

그림 8

와같이 이 變動은 수평선 $\bar{Y}$와 회귀직선과의 수직 거리로서 나타내어진다.

두 變數에 對한 회귀직선에 있어서 a, b는 0이므로 $Y_F=\bar{Y}$가 된다.

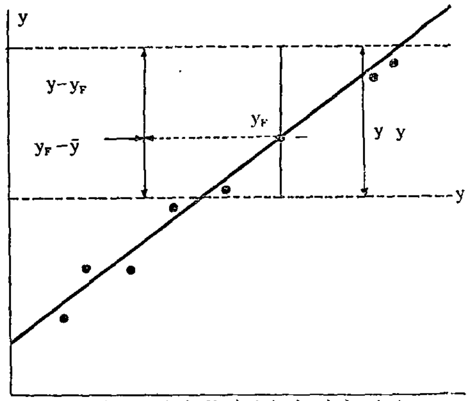


그림 9 2개의 독립변수에 대한 회귀선

그리고 用量間變動의 제곱의 合은 0이다.

이것은 例 2에서 Y_F=117,8 즉 level-line 豫測의 경우와 같이 用量內變動과 全變動은 같음을 나타내고 있다.

豫測方程式 Y_F=117.8+4.7×에서 기울기 b는 全變動의 用量間 變動이 증가하기 때문에 豫測을 개선하여야만 한다.

用量內 變動; 익히 알고 있는 用量內 變動의 제곱의 合은 標準偏差를 計算하는 데서 使用되었다.

S_y는 資料에 對하여 회귀직선의 適合度를 測定하는 것이기 때문에 論理的으로는 適合한 直線에 衣한 用量內變動인 偏差를 기초로 하고 있다.

이러한 用語는 모든 時點이 회귀직선위에 직접적으로 있을 경우에 0가 된다. 따라서 회귀직선은 簡單히 變數 또는 變量의 關係를 정의하는 것으로 一般的인 경우에 있어서 資料는 豫測方程式과 一致하지 않기 때문에 各各의 偏差는 한정(Y)으로부터 회귀직선 까지의 수직거리로서 나타내어 진다.

全變動의 제곱의 合과 用量內變動의 제곱의 合의 比率, $\sum(Y-Y_F)^2/\sum(Y-\bar{Y})^2$는 회귀직선으로 說明되지 않는 全變動의 比率을 측정하는 것이다. 그러므로,

$$1-\frac{(Y-Y_F)^2}{\sum(Y-\bar{Y})^2} \quad (6)$$

는 회귀직선에 依해 說明된 全變動의 比率을 測定한다.

따라서 이式을 決定係數(coefficient of dettermination)라 한다.

相關係數; 決定係數의 제곱근

$$r=\sqrt{1-\frac{\sum(Y-Y_F)^2}{\sum(Y-\bar{Y})^2}} \quad (7)$$

는 一般的으로 相關係數로서 더 잘 알려져 있다.

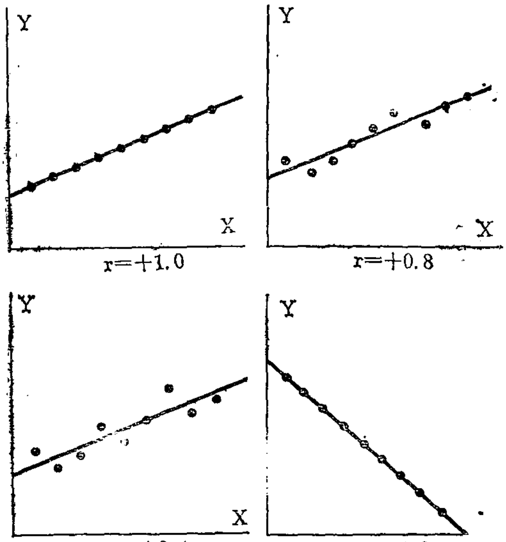
근호안의 價値는 0보다 크거나 1보다 작다.

근호는 둘다 positive와 Negative이기 때문에 r의 價値+1과 −1사이에 있다. (+) 혹은 (−)는 단지 圖表에서 보여주는 것과 같이 회귀직선의 기울기를 나타내는 것이다.

r가 +일때 모든 時點은 위쪽으로 向하는 회귀직선에 놓이게 되고 r가 1과 0사이에 있을 때는 "直線의 양쪽"에 놓이게 된다.

時點이 직선주위에 더 가까이 모이게 되면 될수록 점점 r이 1에 가까와진다.

그림 10을 보면 알 수 있다.



計算 : 充分한 資料가 있을 때 다음 公式에 依하여 r를 計算할 수 있다.

$$r=\frac{N\sum XY-(\sum X)(\sum Y)}{\sqrt{N\sum x^2-(\sum x)^2}\ \sqrt{N\sum Y^2-(\sum Y)^2}} \cdots\cdots(8)$$

이 方程式은 自由度가 充分한 資料로 인하여 고려되지 않을 때 適合한 것이며 資料가 制限되어 있을 때 제곱의 合은 各各 資料와 연관된 自由度에 依하여 나뉘어 진다.

例 2에서 숫자로 나타낸 販賣에 對하여 r를 算이면

$$r=\sqrt{1-\frac{\sum(Y-Y_F)^2/(N-2)}{\sum(Y-\bar{Y})^2/(N-1)}}$$

$$=\sqrt{1-\frac{103.9/3}{224.8/4}}=0.617$$

變動 즉 직회귀선에 依해서 說明되는 變動 Y_F=117.8+4.7×은 실지판매와 평균판매사이의 全變動

의 약 38%(=0.617)² 이다.

信賴度 ; 앞에서 相關係數 0.617을 計算했는데 相關關係에 對하여 資料를 信賴할 수 있는가?

관측회수가 증가하면 信賴度가 높다고 一般的 統計的인 面에서 알 수가 있는 것이다.

그렇지만 95%의 信賴度를 가질려면 관측회수는 얼마程度이어야만 하는가? 이것은 관측회수 N로부터 얻어지는 r 値에 對하여 確率을 나타낸다.

95%, 99%에 對한 N와 r의 決定値가 表 7에 表示되어 있다.

이 表는 事實上 相關係數는 0이 아니라는것을 立證하기 爲하여 試料의 크기 N을 넘어야 한다는 것을 보여주고 있다. 表 6에서 r은 실지적인 0이라는 가정에 依한 統計的 實驗에 依해 開發되었다.

가정은 事實上 앞에서 言及한 바와 같은 r일때 各 信賴度와는 차이가 있다. r가 0이 아니라는 어떤關係를 合理的으로 確證하기에 앞서 r의 計算値는 表에서 보는바와 같이 관측회수 N를 기초로한 決定値를 넘어야만 한다.

試料의 크기가 똑같을 경우 95%에서 99%까지 信賴度를 높일때 r는 95%보다 커야 한다.

<表 6> 95%와 99%의 信賴度에 對한 r의 重要度

N	95%	99%	N	95%	99%	N	95%	99%
10	0.632	0.765	30	0.361	0.463	50	0.279	0.361
12	0.576	0.708	32	0.349	0.449	60	0.254	0.330
14	0.532	0.661	34	0.339	0.436	70	0.235	0.306
16	0.497	0.623	36	0.329	0.424	80	0.220	0.287
18	0.468	0.590	38	0.320	0.413	100	0.413	0.256
20	0.444	0.561	40	0.312	0.413	150	0.161	0.210
22	0.423	0.537	42	0.304	0.393	200	0.139	0.182
24	0.404	0.515	44	0.297	0.384	400	00.98	0.128
26	0.388	0.496	46	0.291	0.376	1.000	0.062	0.081
28	0.374	0.479	48	0.284	0.368			

5. 結　論

지금까지 販賣豫測에 豫測技法과 檢定의 여러 過程을 例를 들어 說明했는데 販賣경쟁이 갈수록 심해지는 現社會의 모든 企業에 있어서는 옛날과 같은 주먹구구식으로서만 예측을 해서는 안된다.

따라서 豫測을 한뒤 그것을 계속 檢定하는 것이 檢定하는 것이 중요하다.

이것은 관리 사이클에서와 같이 계획, 실시, 검사의 過程을 밟는 것이 重要하다고 하겠다.

따라서 豫測만을 하는것은 아무런 意味가 없으므로 豫測에 對한 檢定을 해야한다는 것을 강조해 두고 싶다.

<參考文獻>

1. 統計的 方法〈한국규격협회〉
2. 體系經營學 辭典 徐南源外共著〈文明社〉
3. Modern Production Management by Elwood's Brtta 〈Tnird Edition〉
4. Production System by J.L. Riggs
5. 近代統計學의 理論과 實際 鄭英眞著
6. Hand Book of Industrial Engineering And Management by Ireson Grand(Second Edition)
7. 經營工學便覽 〈日本經營工學會〉