【記事】

# 品質管理에 應用되는 推定과 檢定

株式會社 大成電業 常務 韓 光 植

## 序 言

品質管理를 하다 보면 많은 데이타를 취급하게 되며 이들 데이타를 活用하므로서 情報를 얻을 수 있는 것이다. 特히 샘플링檢査에서는 어느 母集團으로부터 試料를 뽑아 그 結果로 母集團을 推定하게 된다.

品質管理는 工學과 統計學의 結合體라고 볼 수 있다.

合理的이고 科學的인 管理를 實施하는 데는 現在의 狀況을 正確히 把握할 必要가 있으며 計劃(Plan), 實施(Pratice), 監査(Audit), 措置(Action)의 모든 管理段階에서 必要한 것이다

情報는 可能한限 計數化 할 것이 要求되고 計數化되지 않은 情報는 抽象的이어서 解析이나 判斷이 곤란하며 主觀的이고 普遍性이 없고 妥當性이 欠如되여 科學的인 根據가 없게 된다.

데이타는 製品의 狀態, 工程의 狀態 等 여러 狀態를 파악하는데 必要하다. 이 데이타에서 얻은 知識을 統計的으로 處理하고 統計的處理에 따라 다시 工學的으로 檢討함으로서 品質管理는 이루어지는 것이다.

製品이나 半製品에서 試料를 뽑아 試驗하는 것은 단지 그 試料에 대하여 試驗結果를 얻는 것만이 目的이 아니라 그 製品이나 半製品 全體의 性質을 推定하고 檢定하여 生産活動上의 狀態를 알고자 하는 데 있다.

여기에서 간단히 使用되는 推定과檢定을 紹介한다.

## 두 平均値의 推定과 檢定

두 개의 母集團으로부터 取한 샘플의 데이타로부터 두 개의 母平均의 差가 어느 程度인가를 알고 싶을 때 다시 말하면 두 개의 平均値사이에 差異가 있을 때 이들의 試料는 各各 서로 다른 母平均을 갖는 두 母集團으로 부터 뽑혀졌기 때문에 平均値에도 差異가 생긴 것이다.

그렇지 않으면 이 두 試料는 同一한 平均을 갖는 母集團으로부터 Random하게 뽑혀졌지만 各各의 갖고 있는 偶然誤差 때문에 差異가 생긴 것이라 생각할 수 있다.

먼저 平均値의 差의 分布를 생각한다. 平均 $\mu_1$, 標準偏差 $\sigma_1$을 갖는 母集團으로부터 크기 $n_1$인 試料를, random하게 뽑은 試料의 平均 $\overline{x}_1$과 平均 $\mu_2$, 標準偏差 $\sigma_2$인 母集團으로부터, random하게 뽑은 크기 $n_2$인 試料平均 $\overline{x}_2$가 있을 때 이 두 試料가 이루는 差 $\overline{x}_1-\overline{x}_2$의 分布는

平均 ; $\mu_1-\mu_2$

標準偏差 ; $\overline{x}_1-\overline{x}_2=\sqrt{\sigma\overline{x}_1{}^2+\sigma\overline{x}_2{}^2}$

를 갖는 正規分布를 나타낸다.

平均 $\mu$를 갖는 母集團으로 부터 random하게 뽑은 試料의 平均値 $\overline{x}$를 $\mu$의 推定値로 取할

때 $|\overline{x}-\mu|$는 그 推定의 誤差이다. 이 誤差는 試料로 뽑힌 各 個體의 組合의 偶然性에 의하여 생기는 것이므로 人爲的으로 없앨수 없다.

$\overline{x}$에 의하여 $\mu$를 推定할 때 一般的으로 信賴限界로서 하는 것은 이러한 偶然誤差의 必然性을 인정하고 그 區間을 아울러 明示하기 爲한 것이다.

信賴區間은 $\overline{x}-t\alpha\sigma\overline{x}<\mu<\overline{x}+\overline{x}_{\alpha}\sigma_{\bar{x}}$와 같이 되는 데 이는 $|\overline{x}-\mu|<t_{\alpha\bar{x}}$ 程度의 偶然誤差를 不可避한 것으로 認定하는 것이다.

母集團으로 부터 뽑은 試料크기 $n$인 平均値들이 $100(1-\alpha)\%$가 이 區間에 드는 偶然誤差를 가지므로 그 母集團으로 부터 뽑을 수 있는 거의 모든 標本들이 지니는 偶然誤差를 內包하는 區間을 주는 것이다.

그러므로 만약 어떤 試料의 平均値 $\overline{x}$를 얻었을 때 $|\overline{x}-\mu|$의 값이 $|\overline{x}-\mu|<t_{\alpha}\sigma_{\bar{x}}$의 區間에 들면 그 試料는 母平均 $\mu$인 母集團으로부터 random하게 抽出된 한 標本이라는 것을 알게 되며 $|\overline{x}-\mu|\geqq t_{\alpha}\sigma_{\bar{x}}$의 區間에 들면 이 試料는 平均 $\mu$인 母集團으로 부터 뽑혀 나오지 않았다는 것을 알 수 있다.

以上에서 試料의 平均 $\overline{x}$를 보고 그 試料가 어떤 母平均 $\mu$를 갖는 母集團으로 부터 random하게 뽑혀졌는가 아닌가를 다음과 같이 判定한다.

1) 母平均이 $\mu$라고 假定한다.

2) 試料의 平均 $\overline{x}$ 및 標準偏差 $\sigma_{\bar{x}}$를 求한다.

3) $|\overline{x}-\mu|<t_{\alpha}\sigma_{\bar{x}}$ 즉 $\left|\dfrac{\overline{x}-\mu}{\sigma_{\alpha}}\right|<t_{\alpha}$이면 試料의 平均과 母平均의 差는 없다.

4) 一般的으로 $\alpha=0.05$, 및 $\alpha=0.01$이 使用되며 이에 따른 $t_{\alpha}$의 값은 $t$ 分布表에서 0.05의 경우 1.96, 0.01의 경우 2.58을 얻을 수 있다.

有意水準 5% 및 1%인 危險 域은 다음과 같다.

$$\left|\frac{\overline{x}-\mu}{\sigma_{\bar{x}}}\right|\geqq 1.96 (\text{有意水準 } 5\%)$$

$$\left|\frac{\overline{x}-\mu}{\sigma_{\bar{x}}}\right|\geqq 2.58 (\text{有意水準 } 1\%)$$

〔例〕 어느 工場에서 만든 電球의 平均壽命이 1,490時間, 標準偏差 300時間이다.

A機械를 B機械로 바꾸어 100個의 試料를 뽑아 試驗結果 平均壽命 1,500時間이었다.

變更前과 變更後의 차이가 큰가?

위의 문제에서 변경후(1,500時間) 수명이 변경전(1,490時間) 수명보다 크다고 하여 이 變更이 有益한 영향을 주고 있다고 할 수는 없다.

機械變更의 영향의 有無를 判定하는 일은 試料의 平均壽命 $\overline{x}=1,500$時間에 따라 $\mu=1,49$ 時間 $\sigma=300$時間인 假說을 檢定하는 문제가 된다.

試料의 平均壽命 $\overline{x}=1,500$

標準誤差 $\sigma_{\bar{x}}=\dfrac{\sigma}{\sqrt{n}}=\dfrac{300}{\sqrt{100}}=30$

計算 $\left|\dfrac{\overline{x}-\mu}{\sigma_{\bar{x}}}\right|=\dfrac{1,500-1,490}{30}=0.3$

$0.3<t0.05=1.96$

試料의 平均壽命 1,500時間과 母平均 1,490과의 差는 없는 것으로 보고 機械의 變更이 별반 영향을 미치지 않는다고 볼 수 있다.

(例) A, B 두 工場에서 4.0mm의 電線을 購入하려고 한다. A工場 製品의 母集團에서 試料 70個를, B工場 製品의 母集團에서 試料 94個를 random하게 뽑아 치수를 測定한 結果

다음과 같았다. 이 두 工場製品의 차이가 있다고 보는가? (단 과거의 例로 보아 두 공장의 차즉 母平均의 차이는 없다.

| | 시료의 크기 | 平均치수 | 標準偏差 |
|---|---|---|---|
| A 工場 | 70個 | 4.0mm | 0.91mm |
| B 工場 | 94個 | 3.8mm | 1.23mm |

위의 結果에서 4.0>0.38이라고 하여 두 工場의 差異가 있다고 단언할수는 없다. (두 工場의 品質比較는 A工場의 製品이 우수하다는 것은 알 수 있지만 여기에서는 치수의 손실差異를 말함)

1) 두 試料의 母平均의 差가 없다. 즉 $\mu_1-\mu_2=0$이라는 假說를 세운다.

2) 두 試料의 差 $\overline{x}_1-\overline{x}_2$ 및 그의 標準偏差 $\sigma_{\bar{x}1-\bar{x}2}$를 求한다.

3) $\left|\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{x}1-\bar{x}2}}\right|$ ; $\mu_1-\mu_2=0$이므로 $\left|\frac{\overline{x}_1-\overline{x}_2}{\sigma_{\bar{x}1-\bar{x}2}}\right|$를 셈한다.

$\left|\frac{\overline{x}_1-\overline{x}_2}{\sigma_{\bar{x}1-\bar{x}2}}\right| \geqq 1.96=t0.05$ (有意水準 5%; $\alpha=0.05$)이면 두 平均의 差는 심하다.

$\left|\frac{\overline{x}_1-\overline{x}_2}{\sigma_{\bar{x}1-\bar{x}2}}\right| \geqq 2.58=t0.01$ (有意水準 1%; $\alpha=0.01$)이면 두 平均의 差는 매우 심하다.

$\left|\frac{\overline{x}_1-\overline{x}_2}{\sigma_{\bar{x}1-\bar{x}2}}\right| < 1.96$이면 두 平均의 差는 없다고 본다.

4) 檢定結果 差가 없으면 채택되며 두 試料는 다르다고 말할 수 없다.

$$\sigma^2\overline{x}_1=\frac{S_1}{\sqrt{n_1}}=\frac{S_1^2}{n_1}=\frac{0.91^2}{70}$$
$$=0.01183$$
$$\sigma^2\overline{x}_2=\frac{S_2}{\sqrt{n_2}}=\frac{S_2^2}{n_2}=\frac{1.23^2}{94}$$
$$=0.01609$$
$$\therefore\ \sigma_{\bar{x}1-\bar{x}2}=\sqrt{\sigma^2_{\bar{x}1}+\sigma^2_{\bar{x}2}}$$
$$=\sqrt{0.02792}=0.167$$
$$\frac{\overline{x}_1-\overline{x}_2}{\sigma_{\bar{x}1-\bar{x}2}}=\left|\frac{4.0-3.8}{0.167}\right|$$
$$=\frac{0.2}{0.167}\doteqdot 1.2<1.96$$

결과로 보아 差가 없다. 즉 A, B 두 工場의 差는 없다고 볼 수 있다.

## 두 分散의 差의 推定과 檢定

同一한 母分散 $\sigma^2$을 갖는 正規 母集團으로부터 random하게 뽑은 크기 $n_1$과 $n_2$인 두 試料의 分散을 各各 $S_1^2$과 $S_2^2$이라고 하면 두 母分散의 不偏推定値 $\nu_1$, $\nu_2$는

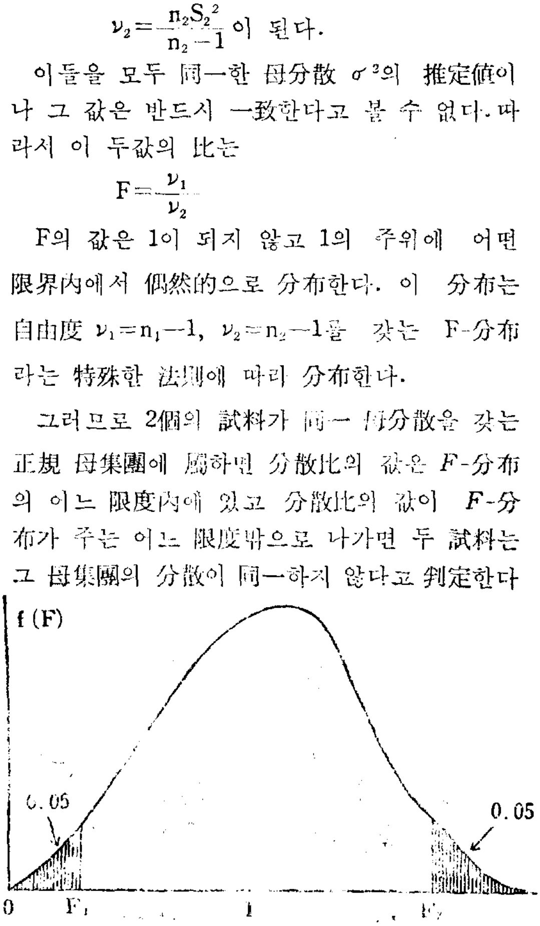
$$\nu_1=\frac{n_1S_1^2}{n_1-1}$$
$$\nu_2=\frac{n_2S_2^2}{n_2-1}$$ 이 된다.

이들을 모두 同一한 母分散 $\sigma^2$의 推定値이나 그 값은 반드시 一致한다고 볼 수 없다. 따라서 이 두값의 比는

$$F=\frac{\nu_1}{\nu_2}$$

F의 값은 1이 되지 않고 1의 주위에 어떤 限界內에서 偶然的으로 分布한다. 이 分布는 自由度 $\nu_1=n_1-1$, $\nu_2=n_2-1$을 갖는 F-分布라는 特殊한 法則에 따라 分布한다.

그러므로 2個의 試料가 同一 母分散을 갖는 正規 母集團에 屬하면 分散比의 값은 $F$-分布의 어느 限度內에 있고 分散比의 값이 $F$-分布가 주는 어느 限度밖으로 나가면 두 試料는 그 母集團의 分散이 同一하지 않다고 判定한다



試料의 크기 $n_1$, $n_2$에 대하여 $F$-分布曲線은 一般的으로 上記 그림과 같이 되며 $F=1$에서 最大로 된다.

만약 有意水準 0.05 (또는 0.01)을 判定의 基準으로 잡으면 두 試料에서 計算한 $F$의 값이

$$F<F_1\text{이거나 } F>F_2\text{이면}$$

이 두 試料分散의 差가 있게 되고 두試料의 母分散은 同一하지 않다고 한다.

分散比 $F$를 셈할 때 $\nu_1$, $\nu_2$中 큰 쪽을 分子에 작은 쪽을 分母에 놓도록 하여 항상 $F>1$과 같이 되도록 하면 差를 判定하는 限界의 값은 $F_2$만으로 된다. 그러므로 이 限界의 값을 주는 $F$-分布表에는 1보다 큰 $F_2$의 값만 주어져 있다.

(例) 다음은 電氣用 更銅線 2.0mm의 引張荷重의 測定値이다. 이 두 日字에 測定한 두 試料의 分散의 差가 있는가 檢定하라.

$x$ ; 8月 5日 測定値

146kg, 143kg, 142kg, 145kg, 140kg

$y$ ; 8月10日 測定値

148kg, 147kg, 149kg, 146kg, 145kg, 147kg

이상의 試料 測定値로 分散의 差의 有意性을 檢定할려면

| No | x | x-μ | (x-μ)² |
| --- | --- | --- | --- |
| 1 | 146 | 2 | 4 |
| 2 | 144 | 0 | 0 |
| 3 | 143 | −1 | 1 |
| 4 | 146 | 2 | 4 |
| 5 | 141 | −3 | 9 |
| 計 | 720 | 0 | 18 |
| 平 均 | 144 | | |

| No | y | y-μ | (y-μ)² |
| --- | --- | --- | --- |
| 1 | 148 | 1 | 1 |
| 2 | 147 | 0 | 0 |
| 3 | 149 | 2 | 4 |
| 4 | 146 | −1 | 1 |
| 5 | 145 | −2 | 4 |
| 6 | 147 | 0 | 0 |
| 計 | 882 | 0 | 10 |
| 平 均 | 147 | | |

x 및 y의 平方合은

$$\sum(x_i-\overline{x})^2=18$$

$$\sum(y_i-\overline{y})^2=10$$

가 된다.

x 및 $y$에 의한 母分散 不偏推定値는

$$\nu x=\frac{18}{5-1}=\frac{18}{4}=4.5$$

$$\nu y=\frac{10}{6-1}=\frac{10}{5}=2.0$$

$$\therefore F=\frac{Vx}{Vy}=\frac{4.5}{2.0}=2.25$$

F의 分子 自由度는 4

F의 分母 自由度는 5 이다.

F分布表에서 $F_5^4(0.05)=5.19$

$F_5^4(0.01)=11.39$

$\therefore F<\frac{F4}{F5}(0.05)$의 結果가 됨으로 이 두試料 分散의 差는 거의 없다고 볼 수 있다.

F 分 布 表

| $V_2$ \ $V_1$ | 4 |
| --- | --- |
| 5 | 5.19<br>11.39 |