行政院國家科學委員會專題研究計劃成果報告

使用個別誤判誤差的雙邊循序篩選程序

Two-Sided Sequential Screening Procedure Based on Individual Misclassified Error

計劃編號: NSC89-2118-M-032-018

執行期限: 89 年 8 月 1 日至 90 年 7 月 31 日

主持人:吳淑妃淡江大學統計系

June 20, 2001

摘要

有些產品的特性不容易檢查、檢查太昂貴或是檢查後會造成毀損，因此不直接觀察表現變數，而藉由觀察數個和表現變數有關的篩選變數，由它們來判定該產品是否被接受，此謂篩選程序。但是為了節省檢驗成本和檢驗時間，多站的循序篩選程序的方法便因應而生。以前大部分的循序篩選程序是針對表現變數具有單邊規格的研究，本研究是針對表現變數有雙邊規格時的循序篩選程序，再以 AOQ 或 TIC 為標準來選出最佳的配置組合。

本篇論文對於k個篩選變數放置於q站的篩選程序，做了變數變換、矩陣運算及運用多變量常態分配的特性發展出一站、兩站和三站的一般性的公式。並且舉了一個將四個篩選變數放置於一站、二站及三站的例子，並利用 Fortran 程式語言計算出每一配置組合的 AOQ 和 TIC，以選出最佳的配置

組合。四站以上的篩選程序，由於軟體的限制，則改採用模擬的方法。本篇論文所使用的篩選變數加權係數是取決於要提高篩選後之產品正確選取率，和Lu(1999)所選擇的加權係數比較的結果，則証實了可降低TIC。

Abstract

A screening procedure which selects items whose performance variable is within a two-sided specification based on observing the correlated screening variables according to the individual misclassification error(IME) is proposed in order to reduce the cost and time effort of inspection. This research is an extended study from the one-sized sequential screening procedure proposed by Wu(1989). The criterion of average outgoing quality(AOQ) and total inspection cost(TIC) are still used to search for the optimal allocation.

The generalized formulas of all desired probabilities for any number of stages of two-sided sequential screening procedures are derived by the use of multivariate normal distribution and matrix operation in this study. An example allocating four screening variables into two-sided SSP, DSP and TSP is illustrated by the computer program via Fortran IMSL using the derived generalized formulas to search for the best allocation.

KEY WORDS AND PHRASES. *Individual nonconforming probability (INP or IME), Average outgoing quality (AOQ), Two-sided sequential screening procedure(SQSP).*

計劃緣由與目的

在這工商業發達的時代，一個產品在出廠前其品質控制愈來愈受重視。良好的品質是每家公司所追求的目標，因為產品品質關係到商譽。所以公司為了永續經營，我們必須嚴格控制出廠產品的品質，而全面檢驗會是一個不錯的方法，但是所花費的成本和時間太多是其缺點，因此循序篩選程序(Sequential Screening Procedure)的方法便因應而生。我們使用多階段篩選程序主要的目的是為了節省檢驗的成本，但必須犧牲一些平均出廠品質Average Outgoing Quality(AOQ)。我們使用了Tsai(1986)所提出的個別誤判機率(Individual Misclassification Error，IME)的理論來發展多階段循序篩選程序，以避免超額允收的問題。

我們考慮表現變數(performance variable)Y是具有常態分配的隨機變數，其平均數和標準差已知。給定表現變數Y為良品的上界U(upper specification)和下界L(lower specification)，即當所有的產品Y值界於U和L之間，我們才接受該產品，反之則拒絕。所謂的篩選程序是不直接測量某項產品的表現變數而測試與其相關的變數，我們稱為篩選變數(screening variable)。

雙邊循序篩選程序的結果與討論

篩選變數係數的選取

根據Tsai(1986)提出的IME的觀念，其循序篩選程序可視為將k個篩選變數配置於r個檢查階段，以減少檢驗時間及成本浪費的設計。假設有k個篩選變數$X_1, X_2, \cdots, X_k$，其與表現變數Y皆具有相關性；假設$X_1, X_2, \cdots, X_k$的檢查成本為$c_1, c_2, \cdots, c_k$。不失一般性，我們假設表現變數Y和k個篩選變數的聯合機率分配為一個多變量常態分配，其期望值為零向量而變異數共變數矩陣為：

$$COV\begin{pmatrix} Y \\ X_1 \\ \vdots \\ X_k \end{pmatrix} = \begin{pmatrix} 1 & \Sigma_{12(1\times k)} \\ \Sigma_{21(k\times 1)} & \Sigma_{22(k\times k)} \end{pmatrix}.$$

當表現變數Y和篩選變數向量X=$(X_1, \cdots, X_k)'$的聯合分配為多元常態分配，期望值為$(\mu_y, \mu_x')'$, $\mu_x' = (EX_1, \cdots, EX_k)$，變異數共變異數矩陣為：

$$\begin{bmatrix} \sigma_y^2 & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix}$$

其中Σ_{yx}和Σ_{xy}為Y和X的共變異數矩陣時，如果令$Z = \beta' X$，其中$\beta = (\beta_1, \beta_2, \cdots, \beta_k)'$，則Y和Z的變異數共變異數矩陣為：

$$\begin{bmatrix} \sigma_y^2 & \beta'\Sigma_{yx} \\ \beta'\Sigma_{xy} & \sigma_z^2 \end{bmatrix}$$

其中$\sigma_z^2 = \beta'\Sigma_{xx}\beta$，Y和Z的相關係數為$\rho_{yz} = \beta'\Sigma_{xy}/\sigma_y\sigma_z$，Tang (1989)提出當$\beta^* = [\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}]^{-1/2}\Sigma_{xx}^{-1}\Sigma_{xy}$時，Y和Z的相關係數有極大值。Tsai(1995)証明表現變數和篩選變數的相關係數會和AOQ成反比，所以和篩選後之正確選取率成正比。為了提高篩選後產品的正確選取率，我們必須使相關係數最大化，也就是將β取為β^*即可。

對於r站循序篩選程序，我們令

$$\begin{aligned} U_1 &= \sum_{i_1 \in A_1} a_{i_1} X_{i_1} \\ U_2 &= \sum_{i_2 \in A_2} a_{i_2} X_{i_2} \\ &\vdots \end{aligned}$$

$$U_r = \sum_{i_r\in A_r} a_{i_r}X_{i_r}$$

其中 $A_1, A_2, \cdots, A_r$ 是集合 $C = \{1,2,\cdots,k\}$ 的r個切割(partitions)，也就是 $A_i \cap A_j = \emptyset, \forall 1 \le i < j \le r$ and $\bigcup_{i=1}^{r} A_i = C$。而篩選變數的r站對應加權係數 $a_{i_1(i_1\in A_1)}, a_{i_2(i_2\in A_2)}, \cdots, a_{i_r(i_r\in A_r)}$ 的值可由下列步驟求得：

步驟1： 對於第一站，重組後可以得到 $\mathrm{Cov}(Y, X_{i_1(i_1\in A_1)})$，我們令為

$$\begin{bmatrix} \Sigma_{yy}^{(1)} & \Sigma_{yx}^{(1)} \\ \Sigma_{xy}^{(1)} & \Sigma_{xx}^{(1)} \end{bmatrix}$$

，為了使產品篩選後的正確選擇率最大化，也就是使相關係數最大化。由Tang(1989)得知 $a_{i_1(i_1\in A_1)}$ 應該給定為 $[\Sigma_{yx}^{(1)}(\Sigma_{xx}^{(1)})^{-1}\Sigma_{xy}^{(1)}]^{-1/2}(\Sigma_{xx}^{(1)})^{-1}\Sigma_{xy}^{(1)}$。

步驟2： 對於第q站 $(2 \le q \le r)$，重組後可以得到 $\mathrm{Cov}(Y, X_{i_q(i_q\in A_q)}, X_{i_{q-1}(i_{q-1}\in A_{q-1})}, \cdots, X_{i_1(i_1\in A_1)})$ 令它等於

$$\begin{bmatrix} \Sigma^{q11} & \Sigma^{q12} \\ \Sigma^{q21} & \Sigma^{q22} \end{bmatrix}$$

，其中 $\Sigma^{q11}=\mathrm{Cov}(Y, X_{i_q(i_q\in A_q)})$ 和 $\Sigma^{q22}=\mathrm{Cov}(X_{i_{q-1}(i_{q-1}\in A_{q-1})}, \cdots, X_{i_1(i_1\in A_1)})$，我們可以求得 $\mathrm{Cov}(Y, X_{i_q(i_q\in A_q)}|X_{i_1(i_1\in A_1)}, \cdots, X_{i_{q-1}(i_{q-1}\in A_{q-1})}) =$ $\Sigma^{q11} - \Sigma^{q12}(\Sigma^{q22})^{-1}\Sigma^{q21}$，我們令其為

$$\begin{bmatrix} \Sigma_{yy}^{(q)} & \Sigma_{yx}^{(q)} \\ \Sigma_{xy}^{(q)} & \Sigma_{xx}^{(q)} \end{bmatrix}$$

，為了使產品篩選後正確選取率最大化，$a_{i_q(i_q\in A_q)}$ 應取為 $[\Sigma_{yx}^{(q)}(\Sigma_{xx}^{(q)})^{-1}\Sigma_{xy}^{(q)}]^{-1/2}(\Sigma_{xx}^{(q)})^{-1}\Sigma_{xy}^{(q)}$。

多變量常態分配 $(Y, U_1, U_2, \cdots, U_r)$ 亦可標準化為多變量標準常態分配 $(Y, V_1, V_2, \cdots, V_r)$。其期望值為零向量，變異數共變異數矩陣為：

$$\Sigma = \begin{pmatrix} 1 & \rho_{01} & \cdots & \rho_{0r} \\ \rho_{10} & 1 & \cdots & \rho_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{r0} & \cdots & \cdots & 1 \end{pmatrix}_{(r+1)\times(r+1)}$$

對於表現變數Y，我們令 $K\gamma_1$ 為其下界，$K\gamma_2$ 為其上界，其中 $P(Y \ge K\gamma_1) = \gamma_1$ 和 $P(Y \le K\gamma_2) = \gamma_2$。亦即，對於一個被檢查項目產品的Y值服從標準常態分配，如果介於 $K\gamma_1$ 和 $K\gamma_2$ 之間，則此項目是適合的，反之則歸類為不適合，所以Y的接受率為 $\gamma_1 + \gamma_2 - 1$。

現在考慮第q站篩選程序，當 $V_1 = v_1, V_2 = v_2, \cdots, V_q = v_q$ 時，此受檢項目的個別誤判誤差定義為 IME_Z $= P(Y \le K\gamma_1, Y \ge K\gamma_2|V_1 =$

$v_1, V_2 = v_2, \cdots, V_q = v_q$)。而選取的比率(Selection ratio)(β)，則定義為選取該項目的比率；拒絕的比率(Rejection ratio)(R)，則定義為拒絕該項目的比率；無法決定的比率(Undecided ratio)(UND)，則定義為無法決定接受或拒絕該項目的比率；Average Outgoing Quality (AOQ)則定義為完成循序篩選程序後，所決定選取項目中，實際為不合適的項目的比率；而Total Inspection Cost(TIC)則定義為當我們在r站中放入k個篩選變數所需的總檢查成本。同時我們定義$emaxa$為接受受檢項目的最大的IME，而$eminr$則為拒絕受檢項目的最小的IME。

在前q站，我們為了計算Z_q，只考慮$(Y, V_1, V_2, \cdots, V_r)$的前q+1個元素。為了求出第q站之IME，我們利用下述之定理。

定理1：*(Morrison(1990))* 如果$(Y, V_1, V_2, \cdots, V_n) = (Y, V^*_{1\times n})$為一多變量常態分配，其期望值為零向量，而變異數共變異數矩陣為：

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, |\Sigma_{22}| > 0$$

則Y在給定$V^* = v^*$的條件下，也是一個常態分配，而其條件期望值為$\Sigma_{12}(\Sigma_{22})^{-1}V^*$，變異數共變異數矩陣為$\Sigma_{11} - \Sigma_{12}(\Sigma_{22})^{-1}\Sigma_{21}$。

由定理1，第q站的IME可化簡如下：

$$Z_q = 1 - \Phi\left(\frac{K\gamma_2 - W_q}{\sqrt{1-\nu^q}}|V_1 = v_1, V_2 = v_2, \cdots, V_q = v_q\right) + \Phi\left(\frac{K\gamma_1 - W_q}{\sqrt{1-\nu^q}}|V_1 = v_1, V_2 = v_2, \cdots, V_q = v_q\right)$$

其中$W_q = b^q \begin{pmatrix} V_1 \\ \vdots \\ V_q \end{pmatrix}, 1 \le q \le r, b^q = \Sigma^q_{12}(\Sigma^q_{22})^{-1}, \nu^q = \Sigma^q_{12}(\Sigma^q_{22})^{-1}\Sigma^q_{21}$ 而$\Phi(\cdot)$是標準常態分配的累積分配函數(c.d.f.)。所以IME $Z_q = 1 - \Phi\left(\frac{K\gamma_2 - W_q}{\sqrt{1-\nu^q}}\right) + \Phi\left(\frac{K\gamma_1 - W_q}{\sqrt{1-\nu^q}}\right)$。為了想了解IME Z_q此函數有何特性，我們有以下之定理。

定理2：IME Z_q的圖形是開口向上，而且是對稱於$W_q = \frac{K\gamma_1 + K\gamma_2}{2}$這一點的拋物線曲線，其最小值會發生在$W_q = \frac{K\gamma_1 + K\gamma_2}{2}$這一點，其值為$2\left[1 - \Phi\left(\frac{K\gamma_2 - K\gamma_1}{2\sqrt{1-\nu^q}}\right)\right]$。

一個檢查項目，在第q站被接受的條件，也就是當

IME $Z_q \le emaxa$

$\Rightarrow 1 - \Phi\left(\frac{K\gamma_2 - W_q}{\sqrt{1-\nu^q}}\right) + \Phi\left(\frac{K\gamma_1 - W_q}{\sqrt{1-\nu^q}}\right) \le emaxa$

時。

當等號成立時，我們令W_q的兩個解為AL_q和AU_q，其中AL_q小於AU_q。也就是當$AL_q \le W_q \le AU_q$，則被檢查項目在第q站會被接受。

一個檢查項目，在第q站被拒絕的條

件，也就是當

IME $Z_q \geq eminr$

$\Rightarrow 1-\Phi\left(\frac{K\gamma_2-W_q}{\sqrt{1-\nu^q}}\right)+\Phi\left(\frac{K\gamma_1-W_q}{\sqrt{1-\nu^q}}\right) \geq eminr$

時。

當等號成立時，我們令W_q的兩個解為RL_q和RU_q，其中RL_q小於RU_q。也就是當$W_q \leq RL_q$或$W_q \geq RU_q$，則被檢查項目在第q站會被拒絕。

一個檢查項目，在第q站無法決定的條件為$RL_q < W_q < AL_q$或$AU_q < W_q < RU_q$。

而最後一站，第r站一定會被決定，其接受的條件為$AL_r \leq W_r \leq AU_r$，反之則拒絕。

為了求取每種配置的AOQ和TIC，我們將利用Wu(1989)之定理，其定理敘述如下。

定理3：假設$(Y, V_1, V_2, \cdots, V_q)$為一個多變量常態分配，其期望值向量為零，變異數共變異數矩陣為

$$\Sigma^q = \begin{pmatrix} 1 & \rho_{01} & \cdots & \rho_{0q} \\ \rho_{10} & 1 & \cdots & \rho_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{q0} & \cdots & \cdots & 1 \end{pmatrix}_{(q+1)\times(q+1)}$$

$$= \begin{pmatrix} 1 & \Sigma_{12}^q \\ \Sigma_{21}^q & \Sigma_{22}^q \end{pmatrix}$$

令$W_i = b^i \begin{pmatrix} V_1 \\ \vdots \\ V_i \end{pmatrix}, 1 \leq i \leq q$，其中$b^i = \Sigma_{12}^i(\Sigma_{22}^i)^{-1}, 1 \leq i \leq q$，

則$(Y, W_q, W_{q-1}, \cdots, W_1)$也是一個多變量常態分配，其期望值為零，變異數共變異數矩陣為

$$COV\begin{pmatrix} Y \\ W_q \\ \vdots \\ W_1 \end{pmatrix} = \begin{pmatrix} 1 & \nu^q & \nu^{q-1} & \cdots & \nu^1 \\ \nu^q & \nu^q & \nu^{q-1} & \cdots & \nu^1 \\ \nu^{q-1} & \nu^{q-1} & \nu^{q-1} & \cdots & \nu^1 \\ \vdots & \vdots & \vdots & \ddots & \nu^1 \\ \nu^1 & \nu^1 & \nu^1 & \cdots & \nu^1 \end{pmatrix}$$

其中$\nu^i = \Sigma_{12}^i(\Sigma_{22}^i)^{-1}\Sigma_{21}^i$，$1 \leq i \leq q$。

經由上述的定理，我們得知一個受檢項目在第q站被接受的機率為：$(1 \leq q \leq r)$

$$\begin{aligned} \beta_q = P(& RL_1 < W_1 < AL_1 \text{ 或 } AU_1 < W_1 < RU_1, \\ & \vdots \\ & RL_{q-1} < W_{q-1} < AL_{q-1} \text{ 或 } AU_{q-1} < W_{q-1} \\ & AL_q \leq W_q \leq AU_q) \end{aligned}$$

同理可得到$R_q, UND_q, q = 1, 2, \cdots, r-1$.

當k個篩選變數被置入r個檢查站時，其總AOQ為：

$$AOQ = \left(\sum_{q=1}^{r} AOQ_q \times \beta_q\right) / \sum_{q=1}^{r} \beta_q,$$

總檢查費用為：

$$TIC = \sum_{i_1 \in A_1} c_{i_1} + \sum_{q=2}^{r} \left(\prod_{m=2}^{q} UND_{m-1} \sum_{i_q \in A_q} c_{i_q} \right)$$

.最後則可求出AOQ and TIC.

例子：一站、二站和三站的比較

由“從原油估計汽油的產量” Lu(1988)的例子中，有四個篩選變數，分別是X_1為原油的重力(AIP)；X_2為原油的水汽壓力(PSIA)；X_3為原油的ASTM 10 percent point(^0F)；X_4為汽油的燃點，Y是原油提練出汽油的比率。假設Y有雙邊允收範圍，其高於允收下界的機率為0.99並且Y低於允收上界的機率為0.95，這樣Y落在被允收的界限內的機率為0.94。再設可接受項目的最大IME為0.05(也就是$emaxa = 0.05$),拒絕項目的最小IME為0.4(也就是$eminr = 0.4$),且X_1, X_2, X_3, X_4的檢查費用分別為100,400,600和1000。假設Y和X_1, X_2, X_3, X_4有多變量常態分配，期望值為零，變異數共變異數矩陣為：

$$COV\begin{pmatrix} Y \\ X_1 \\ \vdots \\ X_k \end{pmatrix} = \begin{pmatrix} 1 & .2463 & .3841 & -.3150 & .702 \\ & 1 & .6206 & -.7002 & -.31 \\ & & 1 & -.9062 & -.30 \\ & & & 1 & .417 \\ & & & & 1 \end{pmatrix}$$

計劃成果自評

我們有6個表,表1,表2表3表示將四個篩選變數配置於一站二站三站時的所有可能情況。配置於三站時,我們用TSP1、TSP2、⋯、TSP36分別來表示這36種組合。而表4、表5、表6分別是四個篩選變數被放置於一、二和三站時，代入公式所求得的各個機率。我們發現其總檢查成本(TIC)會隨著配置站數的增加而減少。所以用循序篩選程序可以降低總檢查成本。當我們分別考慮AOQ和TIC時，我們發現當只有一站時，其AOQ為0.0013，TIC為2100。當篩選變數配置於二站時，如果就品質來考量，我們發現DSP11的AOQ=0.0015為最低，是最佳的組合，但其仍比只有一站時差一點。如果就成本的觀點來考量，我們發現DSP12的TIC=1144.5為最低，是最佳的組合，顯然比一站時省成本。當篩選變數配置於三站時，如果就品質來考量，我們發現TSP29的AOQ=0.0017

為最低，是最佳的組合。如果就成本的觀點來考量，我們發現TSP8的TIC=783.0為最低，是最佳的組合，其顯然比DSP12小。值得注意的是，不管在一站、二站和三站循序篩選程序，產品的總選取率並不會超過表現變數Y的接受率。這個結果應證了在IME理論架構下，可以控制超額允收的問題。所以我們成功的導出雙邊循序篩選程序一站二站三站的所有公式並求出其對應的AOQ, TIC.以求出最佳的配置組合.

參考文獻

Li, L. and Owen, D. B., 1979. Two sided screening procedure in the bivariate case. *Technometrics*, **21**, 79-85.

Lu, D, L., 1999. Simulation Study on Sequential Screening Procedure. *Thesis of Tamkang university.*

Lu, K. H., 1988. Double screening procedure using multiple screening variables. *Thesis of Sun-Yat-sen university.*

Morrison, D. F., 1990. *Multivariate statistical methods.* McGraw-Hill.

Owen, D. B., 1988. Screening by Correlated Variables, *Encyclopedia of Statistical Sciences 8,* edited by S. Kotz and N. L. johnson, John Wiley & Sons, New York, NY, 309-312.

Owen, D. B., Li, L. and Chou, Y. M., 1981. Prediction intervals for screening using a measured correlated variable. *Technometrics*,**23**, 165-170.

Owen, D. B., Mcintire, D. and Seymour, E., 1975. Table using one or two screening variables to increase acceptable product under one-sided specification. *Journal of Quality Technology*, **7**, 127-138.

Tang, K. and Schenider H., 1987. The Effects of Inspection error on a Complete Inspection Plan. *IIE Transactions*, **19**, 421-428.

Tang, J. and Tang, K., 1989. A two-sided screening procedure using several correlated variables. *IIE Transactions*, **21** (4), 333-336.

Taylor, H. C. and Russell, J. T., 1939. The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables. *Journal of Applied Psychology*, **23**, 565-578.

Tsai, H. T., 1986. *Single and Double Screening Procedures Using Indi vidual Misclassification Error*, Ph.D. Thesis, Purdue University.

Tsai, H. T., Moskowitz, H. and Tang, J., 1995. A One-Sided Single Screening Procedure Based on Individual Misclassification Error. *IIE Transactions.* **27**, 695-706.

Wu, S. F., 1989. *Sequential Screening Procedure.* Thesis of Sun-Yat-Sen university.