韓國林學會誌 58:48-59. 1982
Jour. Korean For. Soc. 58:48-53. 1982

判別式에 의한 樹木分類法에 關하여 (I)[1]

—獨逸가문비와 樅榧나무와의 判別分析—

李　光　南[2]

On the Distinction between *Picea koraiensis* Nak. and *Picea abies* (*L.*) Karsten based on the Discriminant Function (I)[1]

Kwang Nam Lee[2]

要　約

現行의 形態學的樹木分類法에 計量的特性을 利用한 所謂 判別分析法에 依할 獨逸가문비(*Picea abies*(L.) Karsten)와 樅榧나무(*Picea koraiensis* Nak.)와의 分類試驗을 實施하고 그 結果를 다음과 같이 要約한다. 1) 本 試驗에서 얻은 判別式과 判別領域은, $Z(x)=Z(x_1, x_2)=0.000379x_1+0.004354x_2-0.311061$ 또는 $Z(x)=Z(x_1, x_2)=0.000379(x_1-60.4428)+0.004354(x_2-66.1851)$, $R_1=\{x \mid 0.000379x_1+0.004354x_2-0.311061\geqq 0\}$, $R_2=\{x \mid 0.000379x_1+0.004354x_2-0.311061<0\}$ 또는 $R_1=\{x \mid 0.000379(x_1-60.4428)+0.004354(x_2-66.1851)\geqq 0\}$, $R_2=\{x \mid 0.000379(x_1-60.4428)+0.004354(x_2-66.1851)<0\}$. 2) 위의 判別領域에 依한 誤判率(誤分類確率)은, $P(2 \mid 1)=P(1 \mid 2)=0.444$로서, $P(2 \mid 1)$와 $P(1 \mid 2)$의 同時誤判率은 $P=44.4\%$. 3) 本 試驗에서 얻은 判別式에 依한 誤判率은 相當히 높게 나타났지만 그의 精度보다는 오히려 判別에 對한 信賴度를 알수 있다는데 보다 큰 意義가 있는 것으로 思料된다.

ABSTRACT

This experiment was carried out to distinguish between *Picea abies (L.)* Karsten and *Picea koraiensis* Nak. by the method of discriminant analysis which is used the metrical continuous characteristic on current morphological plant taxanomy. The results are summarized as follows. 1) The discriminant function and discriminant region from the experiment are $Z(x) = Z(x_1, x_2) = 0.000379x_1 + 0.004354x_2 - 0.311061$ or $Z(x) = Z(x_1,x_2) = 0.000379(x_1-60.442800) + 0.004354(x_2-66.185100)$, $R_1 = \{x \mid 0.000379x_1 + 0.004354x_2 - 0.311061 \geqq 0\}$, $R_2 = \{x \mid 0.000379x_1 + 0.004354x_2 - 0.311061 < 0\}$. 2) The probability of misclassification based on the above discriminant region is $P(2 \mid 1) = P(1 \mid 2) = 0.444$; therefore the probability of simultaneous misclassification of $P(2 \mid 1)$ and $P(1 \mid 2)$ is about 44.4%. 3) The probability of misclassification by the discriminant function resulted from the experiment is recorded as high, but it is thought that there is a considerable meaning to perceive the probability of confidence about the discrimination better than its precision.

Key words : Picea abies; Picea koraiensis; discriminant analysis .

---

[1] 接受 12月 3日 Received December 3, 1982.
[2] 全南大學校 農科大學 College of Agriculture, Chonnam National University, Gwangju, Korea.

緖 言

우리의 周圍에 展開되고 있는 모든 現象 即 우리의 모든 硏究對象은 多面的特性의 結合이라 할 수 있는 것으로 多變量解析法(multivariate analysis)은 이들 많은 現象들이 지니고 있는 多數變量(特性値) 間의 各種關係를 밝히므로써 將來의 豫測이나 變量의 分類, 合成 等을 決定하는 多樣한 技法인 것이다.

最近 高性能 computor의 出現으로 그의 機能이 沈滯되어 있던 多變量解析法은 實際 應用上의 諸種制約條件이 解決됨으로써 生物(測定)學, 心理學, 社會學을 비롯하여 醫學의 疾病診斷에 이르기까지 모든 分野에 關한 高次的인 硏究手段으로서 脚光을 받으며 急進的인 發展을 거듭하고 있다.[1,2,3,4]

위와 같은 發展趨勢에 따라 尙今 거의 應用되고 있지 않는 林學分野의 現實에서 多變量解析法이 劃期的인 硏究手段으로 利用되어 林學硏究의 새로운 紀元이 이루어지길 期待하면서, 著者는 現行 形態學的 植物分類에 判別分析法을 導入適用하여 植物의 各種 特性値의 特徵을 要約, 分類據點을 具體化(形式化)함으로써 보다 科學性이 있는 分類方式이 얻어질것으로 생각하고 今般 *Picea abies* (L.) Karsten 과 *Picea koraiensis* Nak. 와의 判別에 關한 試驗을 實施하고 未洽하나마 그 內容을 報告하는 바이다.

試料 및 方法

1. 試料의 採取

水原市 서울大學校 農科大學構內에서 育成하고 있는 *Picea abies* (L.) Karsten(胸高直徑 ; 25㎝, 樹高 ; 11.1m, 枝下高 ; 2.6m)와 林木育種硏究所構內에서 生育하고 있는 *Picea koraiensis* Nak. (胸高直徑 ; 23.5㎝, 樹高 ; 9.8m, 枝下高 ; 3.7m)를 判別試驗 對象木으로 하여, 兩樹種 다같이 各各의 樹冠中間層의 枝條中에서 4方向의 1年生, 2年生의 枝葉을 各各 1本씩 採取하였다(兩種 다같이 1年生, 2年生 枝葉 各 4本씩 計 8本).

2. 試料의 調製

前項에서 採取된 各 枝葉에 着生한 葉中 摘採以前의 原狀狀態에서의 左右完全 側葉만을 完然하게 남겨두고 其他의 葉은 全部 摘除하여 이것을 白紙上에 原狀으로 놓고 寫眞撮影으로 印畫하였다.

3. 試料의 測定

兩樹種 다같이 各 試料의 寫眞上에서 1年生 枝葉의 左右에서 20個葉, 2年生 枝葉의 左右에서 30個葉씩 樹種當 計 50個葉씩을 規則的으로 指定抽出하여, 抽出된 各 葉身에 對한 2個의 分岐角(中角, 頂角)을 3角定規, divider, vernier-caliper 等을 使用하여 다음과 같은 要領으로 間接測定하였다.[5]

中角(x_2) : 葉身의 基部點과 頂點을 連結한 直線(弦 ; chord)의 50%長을 直線的으로 葉身上에 表示하고, 이 點에서 枝軸에 내린 垂線長을 測定하여 sin x_2를 求하고, 3角函數表에서 x_2의 角度를 索出하였다.

頂角(x_1) : 上述한 弦의 50%長을 100%長(弦의全長)으로 하였을 境遇의 角度.

以上의 分岐角 測定에 있어서 各種길이는 1/50 ㎜, 角度는 1/10度 單位까지 取하였다.

判別分析의 理論(2集團 2變量의 境遇)

1. 資料의 判別能力試驗

本 試驗에서 얻은 特性値(葉身의 分岐角 ; 中角(x_2) 頂角(x_1))에 對한 兩樹種間의 分類據點으로서의 適否 即 判別能力與否를 알아보고자, Bayes' decision law에 依한 條件付確率密度函數를 適用하여[4,8] 判別分析한 結果는 table 1, 2와 같은데, 여기에 適用한 $X=x_i$인 境遇에 있어서의 G_a $(a=1, 2)$의 條件付確率密度函數는,

$$g(a \mid x_i) = \frac{P(a) f_a(x_i)}{P(1) f_1(x_i) + P(2) f_2(x_i)},$$

$(a=1, 2)$, ($P(a)$는 母集團 a의 先驗確率로서 $P(1)+P(2)=1$)……………………(1)

로서 $\dfrac{P(1) f_1(x_i)}{P(1) f_1(x_i) + P(2) f_2(x_i)} \gtrless \dfrac{P(2) f_2(x_i)}{P(1) f_1(x_i) + P(2) f_2(x_i)}$ 에 따라

$P(1) f_1(x_i) \geqq P(2) f_2(x_i)$ 이면 G_1 (*Picea abies* (L.) Karsten)에 歸屬, $P(1) f_1(x_i) < P(2) f_2(x_i)$ 이면 G_2 (*Picea koraiensis* Nak.)에 歸屬되는 것으로 判定 即 標本空間(sample space)을

$$\left.\begin{aligned} R_1 &= \{x_i \mid P(1) f_1(x_i) \geqq P(2) f_2(x_i)\} \\ R_2 &= \{x_i \mid P(1) f_1(x_i) < P(2) f_2(x_i)\} \end{aligned}\right\} \cdots (2)$$

인 2領域으로 分割하였다.

Table 1. Discrimination of group by Bayes' decision law based on x_1.

Division / Class (degree)	G_1 (*Picea abies* (L.) Karsten)		G_2 (*Picea koraiensis* Nak.)		Discrimination of group by Bayes' decision law			
	Frequency	Probability (%)	Frequency	Probability (%)	$g(1 \mid x_i)$	$g(2 \mid x_i)$	Degree belonged to group (G_1 or G_2)	Remarks
40～42			1	1.00		$g(2 \mid 40\sim42)=1.00$	$(40\sim42) \in G_2$	Probability of misclassification $P=P(1)\Sigma_2 f_1(x_i)+P(2)\Sigma'_1 f_2(x_i)$에 依하여 $P=0.05+0.08=0.13$
42～44			1	1.00		$g(2 \mid 42\sim44)=1.00$	$(42\sim44) \in G_2$	
44～46			4	4.00		$g(2 \mid 44\sim46)=1.00$	$(44\sim46) \in G_2$	
46～48			4	4.00		$g(2 \mid 46\sim48)=1.00$	$(46\sim48) \in G_2$	
48～50			8	8.00		$g(2 \mid 48\sim50)=1.00$	$(48\sim50) \in G_2$	
50～52			3	3.00		$g(2 \mid 50\sim52)=1.00$	$(50\sim52) \in G_2$	
52～54			5	5.00		$g(2 \mid 52\sim54)=1.00$	$(52\sim54) \in G_2$	
54～56	1	1.00	3	3.00	$g(1 \mid 54\sim56)=0.25$	$g(2 \mid 54\sim56)=0.75$	$(54\sim56) \in G_2$	
56～58	2	2.00	5	5.00	$g(1 \mid 56\sim58)=0.29$	$g(2 \mid 56\sim58)=0.71$	$(56\sim58) \in G_2$	
58～60	2	2.00	8	8.00	$g(1 \mid 58\sim60)=0.20$	$g(2 \mid 58\sim60)=0.80$	$(58\sim60) \in G_2$	
60～62	6	6.00	3	3.00	$g(1 \mid 60\sim62)=0.67$	$g(2 \mid 60\sim62)=0.33$	$(60\sim62) \in G_1$	
62～64	4	4.00	4	4.00	$g(1 \mid 62\sim64)=0.50$	$g(2 \mid 62\sim64)=0.50$	$(62\sim64) \in G_1$	
64～66	6	6.00	1	1.00	$g(1 \mid 64\sim66)=0.86$	$g(2 \mid 64\sim66)=0.14$	$(64\sim66) \in G_1$	
66～68	9	9.00			$g(1 \mid 66\sim68)=1.00$		$(66\sim68) \in G_1$	
68～70	7	7.00			$g(1 \mid 68\sim70)=1.00$		$(68\sim70) \in G_1$	
70～72	4	4.00			$g(1 \mid 70\sim72)=1.00$		$(70\sim72) \in G_1$	
72～74	3	3.00			$g(1 \mid 72\sim74)=1.00$		$(72\sim74) \in G_1$	
74～76	2	2.00			$g(1 \mid 74\sim76)=1.00$		$(74\sim76) \in G_1$	
76～78	1	1.00			$g(1 \mid 76\sim78)=1.00$		$(76\sim78) \in G_1$	
78～80	1	1.00			$g(1 \mid 78\sim80)=1.00$		$(78\sim80) \in G_1$	
80～82	2	2.00			$g(1 \mid 80\sim82)=1.00$		$(80\sim82) \in G_1$	
Total	50	50.00	50	50.00				

Table 2. Discrimination of group by Bayes' decision law based on x_2.

Division / Class (degree)	G_1 (*Picea abies* (L.) Karsten)		G_2 (*Picea koraiensis* Nak.)		Discrimination of group by Bayes' decision law			
	Frequency	Probability (%)	Frequency	Probability (%)	$g(1 \mid x_i)$	$g(2 \mid x_i)$	Degree belonged to group (G_1 or G_2)	Remarks
42～44			1	1.00		$g(2 \mid 42\sim44)=1.00$	$(42\sim44) \in G_2$	Probability of misclassification $P=P(1)\Sigma_2 f_1(x_i)+P(2)\Sigma'_1 f_2(x_i)$에 依하여 $P=0.02+0.05=0.07$
44～46			1	1.00		$g(2 \mid 44\sim46)=1.00$	$(44\sim46) \in G_2$	
46～48			2	2.00		$g(2 \mid 46\sim48)=1.00$	$(46\sim48) \in G_2$	
48～50			4	4.00		$g(2 \mid 48\sim50)=1.00$	$(48\sim50) \in G_2$	
50～52			4	4.00		$g(2 \mid 50\sim52)=1.00$	$(50\sim52) \in G_2$	
52～54			2	2.00		$g(2 \mid 52\sim54)=1.00$	$(52\sim54) \in G_2$	
54～56			4	4.00		$g(2 \mid 54\sim56)=1.00$	$(54\sim56) \in G_2$	
56～58			5	5.00		$g(2 \mid 56\sim58)=1.00$	$(56\sim58) \in G_2$	
58～60			5	5.00		$g(2 \mid 58\sim60)=1.00$	$(58\sim60) \in G_2$	
60～62			5	5.00		$g(2 \mid 60\sim62)=1.00$	$(60\sim62) \in G_2$	
62～64			6	6.00		$g(2 \mid 62\sim64)=1.00$	$(62\sim64) \in G_2$	
64～66	2	2.00	6	6.00	$g(1 \mid 64\sim66)=0.33$	$g(2 \mid 64\sim66)=0.67$	$(64\sim66) \in G_2$	
66～68	6	6.00	2	2.00	$g(1 \mid 66\sim68)=0.67$	$g(2 \mid 66\sim68)=0.33$	$(66\sim68) \in G_1$	
68～70	5	5.00	2	2.00	$g(1 \mid 68\sim70)=0.71$	$g(2 \mid 68\sim70)=0.29$	$(68\sim70) \in G_1$	
70～72	2	2.00	1	1.00	$g(1 \mid 70\sim72)=0.67$	$g(2 \mid 70\sim72)=0.33$	$(70\sim72) \in G_1$	
72～74	9	9.00			$g(1 \mid 72\sim74)=1.00$		$(72\sim74) \in G_1$	
74～76	8	8.00			$g(1 \mid 74\sim76)=1.00$		$(74\sim76) \in G_1$	
76～78	4	4.00			$g(1 \mid 76\sim78)=1.00$		$(76\sim78) \in G_1$	
78～80	4	4.00			$g(1 \mid 78\sim80)=1.00$		$(78\sim80) \in G_1$	
80～82	5	5.00			$g(1 \mid 80\sim82)=1.00$		$(80\sim82) \in G_1$	
82～84	4	4.00			$g(1 \mid 82\sim84)=1.00$		$(82\sim84) \in G_1$	
84～86	1	1.00			$g(1 \mid 84\sim86)=1.00$		$(84\sim86) \in G_1$	
Total	50	50.00	50	50.00				

table 1, 2의 內容 및 fig. 1에 따라 本 試驗에서 判別特性으로 使用하게된 2分岐角(x_2, x_1)은 다같이 兩樹種間의 有力한 判別(分類) 因子로서 採擇할 수 있음이 確認되었다.

2. 判別函數의 誘導

2個의 母集團 G_1(*Picea abies* (L.) Karsten), G_2(*Picea koraiensis* Nak.)는 다같이 2變量(x_1, x_2)의 正規分布 $N(\mu_1, \Sigma)$, $N(\mu_2, \Sigma)$를 이루게될 것이 分明하므로 이들 2母集團을 各各 正規母集團이란 前提下에서 2分岐角 x_1, x_2에 關한 判別函數式의 推定이 可能하게된다. 이 때 Σ는 dispersion matrix로서 兩母集團의 共通인 것으로 假定한다.

따라서 2變量正規分布의 確率密度函數(probability density function)은,

$$f(x)=f(x_1, x_2)=\frac{1}{2\pi|\Sigma|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)\right\}$$

에 依해서 定義되므로, 2母集團에 依해서 이루어진 標本空間이 誤判率이 最小가 되게 2個領域(R_1, R_2)으로 分割될 수 있도록 하는 境界의 函數를 얻기 위하여, 母集團 G_1, G_2의 各各의 確率密度函數 $f_1(x)$, $f_2(x)$의 比를 求함으로써 線形判別函數(linear discriminant function)가 얻어진다.[4,9]

即 $$\frac{f_1(x)}{f_2(x)}=\frac{1/2\pi\sqrt{|\Sigma|}\ \exp\{-\frac{1}{2}(x-\mu_1)'\Sigma^{-1}(x-\mu_1)\}}{1/2\pi\sqrt{|\Sigma|}\ \exp\{-\frac{1}{2}(x-\mu_2)'\Sigma^{-1}(x-\mu_2)\}}=\exp\{x'\Sigma^{-1}(\mu_1-\mu_2)-\frac{1}{2}(\mu_1+\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)\}\cdots(3)$$

에 따라서 線形判別函數(Z)

$$Z=x'\Sigma^{-1}(\mu_1-\mu_2)-\frac{1}{2}(\mu_1+\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)\quad\cdots\cdots\cdots\cdots(4)$$

를 얻게된다.

上記 判別函數式에 따라 標本空間을 R_1, R_2로 分割하려면 分割의 基準點(限界點)이 決定되어야 하는데 이 限界點으로서는, $x_i \in G_1$이 $x_i \in G_2$로, $x_i \in G_2$가 $x_i \in G_1$로 誤判되는 境遇에 發生하는 損失을 各各 $C(2|1)$, $C(1|2)$라 하고, 前節에서와 같이 x_i가 G_1에서 選拔될 수 있는 先驗確率(mathematical probability)을 $P(1)$, G_2에서 選拔될 수 있는 先驗確率을 $P(2)$라고 할 때, 그의 比値인

$$k=\frac{P(2)\,C(1|2)}{P(1)\,C(2|1)}$$ 를 使用하면 된다.[4,9]

그런데 判別函數式(4)는 兩母集團(G_1, G_2)의 比로서 이것은 exp函數의 指數일 뿐만 아니라 exp函數는 單調增加函數(monotone increasing function)이므로 結局 R_1, R_2의 領域은 다음에 依해서 劃定된다.

$$\left.\begin{aligned}R_1&=\{x\,|\,x'\Sigma^{-1}(\mu_1-\mu_2)-\tfrac{1}{2}(\mu_1+\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)\geqq\log k\}\\R_2&=\{x\,|\,x'\Sigma^{-1}(\mu_1-\mu_2)-\tfrac{1}{2}(\mu_1+\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)<\log k\}\end{aligned}\right\}\cdots(5)$$

3. 誤判(確)率의 誘導

위에서 推定된 判別函數(4)에 있어서의 變量 x(x_1, x_2)가 $N(\mu_1, \Sigma)$ 또는 $N(\mu_2, \Sigma)$의 어느것을 이루기 때문에, 그의 線形結合(linear combination)인 Z(4)도 또한 어느 1變量正規分布를 이루게된다. 前節에서 劃定된 判別領域(5)에 依해서 判別을 行하였을 境遇에 일어나는 誤判(確)率을 求하기 爲하여

1) x가 G_1에 屬한 것이라 假定하였을 境遇의 Z(4)의 平均($E_1\{Z\}$)과 分散($V_1\{Z\}$),

2) x가 G_2에 屬한 것이라 假定하였을 境遇의 Z(4)의 平均($E_2\{Z\}$)과 分散($V_2\{Z\}$)를 算出하면, 各各

$$\left.\begin{aligned}E_1\{Z\}&=\tfrac{1}{2}(\mu_1-\mu_2)'\Sigma^{-1}(\mu_1-\mu_2),\\V_1\{Z\}&=(\mu_1-\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)\\E_2\{Z\}&=-\tfrac{1}{2}(\mu_1-\mu_2)'\Sigma^{-1}(\mu_1-\mu_2),\\V_2\{Z\}&=V_1\{Z\}\end{aligned}\right\}\cdots(6)$$

로 된다.

이 때 $D^2=(\mu_1-\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)\ \cdots\cdots\cdots\cdots(7)$

라 하면, Z는 前者인 境遇에 있어서는 $N(\frac{D^2}{2}, D^2)$,

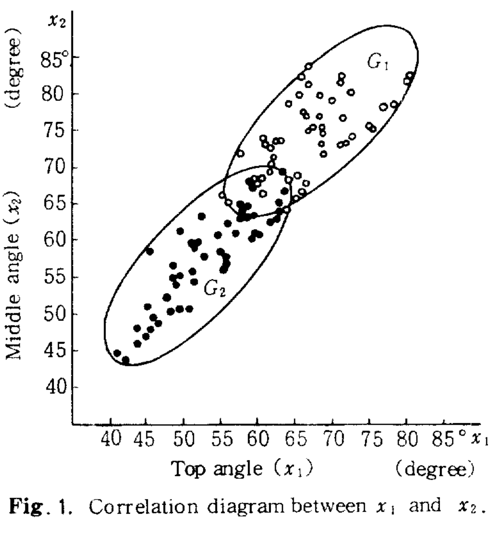

Fig. 1. Correlation diagram between x_1 and x_2.

後者인 境遇에 있어서는 $N(-\frac{D^2}{2},\ D^2)$인 分布를 이루게됨을 알 수 있다.

이 때 D^2는 Mahalanobis' distance라 稱하는 것으로서 判別良否의 程度를 表示하는 尺度가 된다.[4,7]

따라서 判別領域(5)에 依한 判別에서 생기는 誤判(確)率 即 $P(2|1)$ ($x_i \in G_1$이 $x_i \in G_2$로 誤判되는 確率)와 $P(1|2)$ ($x_i \in G_2$가 $x_i \in G_1$로 誤判되는 確率)는 標準正規分布函數(standard normal distribution function)에서의 確率推定理論에 依據 다음과 같이 推定된다.

$$P(2|1)=\int_{-\infty}^{l}\frac{1}{D\sqrt{2\pi}}\exp\left\{-(t-\frac{D^2}{2})^2/2D^2\right\}dt$$
$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(l-\frac{D^2}{2})/\sqrt{D^2}}\exp(-\frac{y^2}{2})dy$$
$$\left(y^2=\frac{(t-\frac{D^2}{2})^2}{D^2}\text{ 로 變數變形}\right)$$
$$P(1|2)=\int_{l}^{\infty}\frac{1}{D\sqrt{2\pi}}\exp\left\{-(t+\frac{D^2}{2})^2/2D^2\right\}dt$$
$$=\frac{1}{\sqrt{2\pi}}\int_{(l+\frac{D^2}{2})/\sqrt{D^2}}^{\infty}\exp(-\frac{y^2}{2})dy$$
$$\left(y^2=\frac{(t+\frac{D^2}{2})^2}{D^2}\text{ 로 變數變形}\right)\quad\cdots\cdots(8)$$

判別式 및 誤判(確)率의 推定

1節의 資料에 對한 判別能力 試驗에서 判別有力因子로 認定된 變量(分岐角 ; 中角(x_2), 頂角(x_1))에 關한 判別式과 이에 依한 誤判(確)率을 다음과 같이 推定하였다. (Fig. 1)

table 3과 같은 兩樹種 *Picea abies* (L.) Karsten (G_1)와 *Picea koraiensis* Nak. (G_2)에 對한 各種 統計定數와 兩樹種 全 data에 對한 dispersion matrix Σ(variance covariance matrix) 및 그의 inverse matrix Σ^{-1}

即 $\Sigma=\begin{bmatrix}3559.365111 & 2768.898402\\ 2768.898402 & 3587.360490\end{bmatrix}$,

$\Sigma^{-1}=\begin{bmatrix}0.000703 & -0.000543\\ -0.000543 & 0.000698\end{bmatrix}$를 求하게 됨에 따라 Fisher's discriminant function 인

$\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}0.000703 & -0.000543\\ -0.000543 & 0.000698\end{bmatrix}\begin{bmatrix}13.4220\\ 16.6794\end{bmatrix}$와

$\frac{1}{2}\begin{bmatrix}120.8856 & 132.3702\end{bmatrix}\begin{bmatrix}0.000703 & -0.000543\\ -0.000543 & 0.000698\end{bmatrix}\begin{bmatrix}13.4220\\ 16.6794\end{bmatrix}$를 풀어 整理함으로써 다음 判別式을 얻게 되었다.

$$Z(x)=Z(x_1,\ x_2)=0.000379x_1+0.004354x_2-0.311061$$

또는 $$Z(x)=Z(x_1,\ x_2)=0.000379(x_1-60.442800)+0.004354(x_2-66.185100)\cdots\cdots(9)$$

本 試驗의 境遇 G_1, G_2가 $N(\mu_1,\ \Sigma)$, $N(\mu_2,\ \Sigma)$라는 前提下에서는 前述한 바의 誤判으로 因한 損失 $C(2|1)$와 $C(1|2)$를 各各 1 即 $\{C(2|1)=C(1|2)=1\}$로, $P(1)$과 $P(2)$를 서로 같은 $\frac{1}{2}$ 即 $P(1)=P(2)=\frac{1}{2}$로 假定하여도 支障이 없으므로 兩領域의 分割限界點은 $l=\log k=\log 1=0$ 으로 된다.[4)]

따라서 兩母集團 G_1, G_2의 判別領域(R_1, R_2)은,

$$R_1=\{x\,|\,0.000379x_1+0.004354x_2-0.311061\geqq 0\}$$
$$R_2=\{x\,|\,0.000379x_1+0.004354x_2-0.311061<0\}$$

또는

$$R_1=\{x\,|\,0.000379(x_1-60.4428)+0.004354(x_2-66.1851)\geqq 0\}$$
$$R_2=\{x\,|\,0.000379(x_1-60.4428)+0.004354(x_2-66.1851)<0\}\quad\cdots\cdots\cdots\cdots(10)$$

와 같이 劃定된다.

推定된 判別式에 依해서 本 試驗에 使用한 兩樹種(G_1, G_2) 別의 各各의 原 data (x_1, x_2)에 對한 判別函數値를 計算하여, $Z(x_1,\ x_2)\gtrless 0$에 따라 集團判別하여보면 table 4와 같다.

table 4에 依하면 原所屬 G_1인 것이 G_2로, G_2인 것이 G_1로 誤判(誤分類)된 것이 各各 3個, 4個 即 3/50, 4/50로서 各各의 實際誤判率이 매우 낮

Table 3. Several statistical constants of 2 leaf-angles.

Variate	Mean of	Mean of	Difference	Sum	Coefficient of correlation (ρ)	
(x)	G_1 (μ_1)	G_2 (μ_2)	($\mu_1-\mu_2$)	($\mu_1+\mu_2$)	G_1	G_2
x_1	67.1538	53.7318	13.4220	120.8856	0.624	0.887
x_2	74.5248	57.8454	16.6794	132.3702	($\bar{\rho}$ = 0.756)	

은 6%, 8%로 나타나므로써 本 硏究에서 얻은 判別式은 相當히 높은 實用性을 지닌것으로 推斷된다. (table 4 省略)

또한 위에서 劃定된 判別領域(10)에 依한 分類(判別)에서 誤判될 수 있는 確率(誤判率)은, Mahalanobis' distance D^2가

$$(13.4220 \quad 16.6794)\begin{bmatrix} 0.000703 & -0.000543 \\ -0.000543 & 0.000698 \end{bmatrix}\begin{bmatrix} 13.4220 \\ 16.6794 \end{bmatrix}$$

의 計算에 따라 $D^2 = 0.077709$, 또한 前述한 바의 $l = 0$에 따라

$$\left(l - \frac{D^2}{2}\right) \Big/ \sqrt{D^2} = 0.139384$$ 가 되므로

$$\left.\begin{aligned} P(2 \mid 1) &= \int_{-\infty}^{-0.139384} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = 0.444 \\ P(1 \mid 2) &= \int_{0.139384}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = 0.444 \end{aligned}\right\} \cdots$$

....(11)이 된다.

따라서 $P(2 \mid 1)$와 $P(1 \mid 2)$의 同時誤判(確)率은 兩者의 半分의 合인 44.4%가 된다.

위의 結果에 따르면 誤判(確)率이 바라는것 보다는 높은 것으로 나타났는데 이것은 兩樹種 다같이 變量間의 共線性(colinearity)에 起因된 것으로 思料되므로 變數選擇을 通한 보다 깊은 試驗硏究가 있어야 할것으로 생각된다.[7]

引 用 文 獻

1. 近藤正已. 1961. 判別函數による生長豫測の硏究(豫報). 日林誌43(1): 1 − 6.
2. 編集部. 1956. 判別函數の問題と解き方(1). 氣象と統計7(1): 20 − 22.
3. 編集部. 1956. 判別函數の問題と解き方(2). 氣象と統計7(3): 84 − 86.
4. 河口至商. 1980. 多變量解析入門. 森北出版社: 79 − 91.
5. 木梨謙吉, 宮島 寬. 1965. Discriminant Functionによるスギ品種の分類法について. 日本林學會 九州支部大會講演集 19: 105−106.
6. Morrison, D. F. Multivariate statistical methods. McGraw-Hill: 230 − 246.
7. 奧野忠一, 久米 均, 芳賀敏郎, 吉澤正. 1980. 多變量解析法. 日科技連出版社: 128 − 131, 259 − 321.
8. 武藤眞介. 1976. 多變量解析入門. 現代數學社: 57 − 64.
9. 竹內 啓, 柳井晴夫. 1980. 多變量解析の基礎. 東洋經濟新報社: 189 − 196.