

농업실험에서 直線分離函數의 利用

蔡 永 岩

문헌농공대학

Linear Discriminant Analysis in Agricultural Experiment.

Young Am Chae

Technische Universität München, West Germany

ABSTRACT

Using head length and head width of two wheat monosomic lines linear discriminant function of these two variables was calculated and also illustrated how one can effectively classify unknown individuals into a correct group belonging by means of this linear discriminant function in reverse. Brief suggestion on the utilization of this analysis in genetics and breeding program was given.

어느 群을 다른 群으로부터 分離判別하러 할때 直線分離函數가 利用되고 있는데 植物分類學에서는 種이나 變種을 다른 變種과 區別하러 할때 쓰이고 社會學에서는 雇傭主가 雇傭人을 適材適所에 配置하러 하거나 潛在行政能力을 把握하러 할때 或은 飛行訓練適性을 檢定하는데 이 直線分離函數가 쓰여오고 있었다.

이와같은 未分類의 個體들을 어떤 群으로 分類코저 하는 問題는 農學에서도 漸次로 그 要求度가 높아 가고 있다. 예를들면 어떤 地域에서 適應하는 品種을 選定한다면가 育種材料母本을 効率的으로 選定하는등 또는 病理學分野에서 새로운 病原性群系로 判定하는 實驗設計等に 應用될 수 있을 것이다.

本報告는 이 直線分離函數를 小麥의 Genome에 따른 草型의 分離判別에의 利用可能性 與否를 檢討한 것이다.

本人에게 이 研究를 할수 있도록 材料와 研究의 便宜를 提供해준 本大學當局에 感謝한다.

材料 및 方法

本大學에서 育成中인 Caribo monosomic (Dr. E.R.

Sears에게서 分讓받은 Chinese spring monosomics에 다 Caribo品種을 戻交雜하여 만든 monosomics series) 5A (chromosome IX)와 2D (chromosome XX)의 穗型의 差異가 棼에 着眼하여 穗長과 穗幅 어느 形質이 어떻게 穗型決定에 寄與하는 가를 檢討한바 있는데¹⁾ 本研究에서는 前報에서 利用한 實測值를 그대로 利用하였다. 따라서 材料 및 育成方法은 前報「多變量同時分析法에 依한 小麥 monosomic 5A와 2D의 穗型分析」을 參照해 주기 바란다.

結果 및 考察

2개 monosomic의 穗長과 穗幅을 表1에 表示하였는데 다른 數値의 誘導는 前報 多變量同時分析法에서 說明한바와 같다.

Table 1. Head length and head width of two Caribo monosomics 5A and 2D.

5A		2D	
(1) Y_{1j} HL	(2) Y_{1j} HW	(1) Y_{2j} HL	(2) Y_{2j} HW
11.1	1.2	9.2	1.4
11.2	1.1	11.8	1.3
12.0	1.2	10.0	1.4
13.7	1.3	10.7	1.4
13.2	1.1	10.6	1.4
12.3	1.2	9.5	1.3

monosomic 5A의 2가지 觀測值 $Y_{1j}^{(1)}$, $Y_{1j}^{(2)}$ 를 scale-觀測으로 代替하면

$$L_{1j} = a_1 Y_{1j}^{(1)} + a_2 Y_{1j}^{(2)}$$

로 되는데 여기서 j 는 조사개체수이며 a_1 과 a_2 는 항수이다. 같은 例로 2D에서는

$$L_{2j} = a_1 Y_{2j}^{(1)} + a_2 Y_{2j}^{(2)}$$

가 된다.

그러면 이 2식의 선형복합인

$$L = a_1 Y^{(1)} + a_2 Y^{(2)}$$

는 직선분리函數가 된다.

여기서 우리는 a_1 과 a_2 를 알게되면 L 值를 求할수 있게 될 것이다.

즉 $\underline{a} = \underline{d} \underline{s}^{-1}$ 은 일반식으로 여기서는

$$\underline{a} = (d_1 \ d_2) \underline{s}^{-1} = (1.45 \ -0.183) \begin{pmatrix} 0.6104 & -2.7248 \\ -2.7248 & 190.7357 \end{pmatrix}$$

$$= (1.6889 \ -40.2180)$$

으로 計算되며 다시 이로부터

$$a_1 = 1.6889 \text{ 그리고 } a_2 = -40.2180$$

이 됨을 알수 있다. 따라서 이 경우 직선분리函數는

$$L = 1.6889 Y^{(1)} - 40.2180 Y^{(2)}$$

가 된다.

이 직선분리函數를 利用하여 次後調査되는 個體들을 區別하여 monosomic 5A나 2D어디에 속하는가를 決定할 수 있게 될것이다. 다시말하면 어느個體의 穗長과 穗幅을 알면 이 L 值를 求하여 5A나 2D 어느 것에 屬하는가를 알 수 있게 된다.

本 實驗의 경우에는 $L - aY_2'$ 值가

$$D^2/2 = 5.3267$$

보다 큰 경우에는 5A에 屬하고 작은 경우에는 2D에 屬하게 된다.

좀더 具體的으로 이 分離函數의 分離能力을 分明히 하기 爲하여 5A와 2D에 屬한 各個體의 穗長과 穗幅測定值를 使用하여 該當個體의 直線分離函數值를 計算하여 表2에 表示하였다.

計算은 $L_{11} = 1.69 \times 11.1 - 40.22 \times 1.2 = -29.50$ 으로

Table 2. Values of the linear discriminant function computed from table.

	5A	2D
1	- 29.50	- 40.76
2	- 25.31	- 32.35
3	- 27.98	- 39.41
4	- 29.14	- 38.23
5	- 21.93	- 38.40
6	- 27.47	- 36.23
Total	-161.33	-225.38
Mean	- 26.89	- 37.56

로 以下 이와같이 하여 計算되었다.

表2에서 보면 5A와 2D에 屬한 各個體의 函數值는 조금도 重複되지 않고 있다. 結局은 穗長과 穗幅을 同時에 考慮하여 分析하므로써 各個體를 正確하게 該當群에 按配시킬 수 있게 된다. 이러한 結果는 單的으로 이 直線分離函數의 機能을 表現하는 것이라고 말할 수 있을 것 같다.

지금 2D의 1번個體를 5A와 2D의 어느群에 屬하는가를 決定하기 위하여 이삭길이 9.2와 이삭폭 1.4를 取하여 直線分離函數를 求하면

$$L_{22} = -40.76 \text{ 이 된다. 이 값을}$$

$$\bar{L}_2 = -37.56 \text{ 과 比較하면}$$

$$L_{22} - \bar{L}_2 = -40.76 + 37.56 = -3.20 \text{으로}$$

이 -3.2 는 $D^2/2 = 5.33$ 보다 작으므로 2D群에 屬하게 된다고 볼수 있다.

이 直線分離函數의 意味를 더 分明히 理解하기 爲하여 이 函數와 觀測值間의 相關關係를 다음 一般式

$$\hat{\rho}(Y^{(k)} \ L) = d_i / \sqrt{s_{kk} D^2}$$

에서 求하여 본다.

먼저 이삭길이와 이 直線分離函數와의 相關은

$$\hat{\rho}(Y^{(1)} \ L) = 1.95 / \sqrt{1.75 \times 10.6533} = 0.451$$

이삭폭과의 相關은

$$\hat{\rho}(Y^{(2)} \ L) = -0.183 / \sqrt{0.0056 \times 10.6533} = -0.750$$

으로 길이와의 相關은 폭과의 相關보다 낮게 나타났다. 이것은 前報 多變量同時分析法에서 指摘한 바와 같이 길이에 對한 信賴區間이 零을 包含하지 않아 穗長은 穗型決定에 直接的으로 크게 影響하지 않았음을 堪案할때 더 理解가 잘 된다. 결국 a_1 과 a_2 는 각각 穗長과 穗幅에 對한 區別係數가 되며 a_2 가 a_1 보다 區別能力이 컸다고 말할 수 있다.

이 相關關係를 좀더 擴張해 보면 많은 形質 或은 變量을 取扱할 때 比較的 容易하게 個體를 區別하기 爲하여 調査하는 形質數를 最少限으로 줄일 수도 있을 것이다. 즉 直線分離函數와 相關이 낮은 形質은 考慮에서 除外한 다음 相關이 높은 形質들만 가지고 다시 새로이 直線分離函數를 計算하면 보다 少數의 形質個體들을 効率的으로 判別할 수 있을 것이다.

摘 要

直線分離函數를 利用하여 小麥 Caribo monosomics 5A와 2D의 穗型을 比較하므로써 直線分離函數의 農業에서의 利用性을 檢討하였다.

直線分離函數를 計算하여 2가지 穗型을 完全히 分

離할 수 있었으며 그 결과는 前報 多變量同時分析法의 결과와 一致하였다.

直線分離函數의 意味의 擴大解釋으로 調査形質數를 減少시킬 수 있는 可能性을 검토 하였다.

引 用 文 獻

1. 蔡永岩 1976, 多變量同時分析法에 依한 小麥 monosomics 5A와 2D의 穗型分析. 育種지 제8권 제2호 ; 112—114.
2. KRAMER, C.C. 1972, A first course in methods of multivariate analysis. pp. 351. Virginia Polytechnique Institute and State University Press.

SUMMARY

The application of linear discriminant analysis was discussed utilizing the data on the panicle type of Caribo monosomics 5A and 2D.

A numerical example computed from linear discriminant function showed distinct two different panicle types, and this coincides well with the one obtained by multivariate analysis.

The feasibility in which the number of traits will be reduced by the extended implications of linear discriminant analysis was also discussed.