

群集檢測方法在植物病理學上之應用

蔣國司^{1,2} 賴信宏¹

¹ 台中市 國立中興大學農藝學系

² 聯絡作者，電子郵件：kucst@dragon.nchu.edu.tw；傳真：+886-4-2285-6497

接受日期：中華民國 99 年 8 月 28 日

摘要

蔣國司、賴信宏. 2010. 群集檢測方法在植物病理學上之應用. 植病會刊. 19: 117-126.

群集檢測 (group testing) 是將單一檢測單位結合成群集，並以群集當做檢測單位來進行測試。當被檢測對象陽性機率低且檢測成本所費不貲時，這種檢測方法特別有其價值。此方法在植物病理界受到相當大的重視，主要始於 Swallow 在 1985 年發表於 *Phytopathology* 之報告，從此之後探討此方法之理論與應用相繼在植物病理界期刊出現。這些研究指出在不增加成本與勞力下，群集檢測對於取樣過程中所得資訊能提供相當大的改進，因此台灣植物病理學家熟悉群集檢測之方法實有其必要性。本文首先解釋說明群集檢測之原理及程序，進而再以台灣植物病理界三個實例來說明其應用性，最後並討論群集檢測未來發展之可能方向。

關鍵詞：群集檢測、估計、分類、均方誤差、成本

緒言

群集檢測最早是在 1943 年由 Dorfman 在統計期刊所提出應用研究於醫學之方法，目的為當感染率相當低以致於檢測單一個體缺乏效率的情況下，在調查群集中發展一套有效率挑出有病個體之方法⁽³⁾。在植物病理界最有名的例子，首推 Hughes 與 Gottwald 為了有效決定南美立枯病 (*Citrus tristeza virus*) 之發生率 (incidence) 所做的研究⁽⁶⁾。由於此病害對於柑橘產量有極大威脅，美國農部需以有效率方式得知植株位置，進而加以剷除。在此研究中使用四株為一單位之群集進行檢測，而此取樣方法現今在美國與其他地方已廣泛被使用，在不增加成本與時間下，大大提高了所得資訊之精確度。

台灣植物病理學者雖經常用到群集檢測的概念，但以群集檢測為核心來探討的應用實例幾乎沒有，故以下內容除將介紹群集檢測的原理與方法外，亦將以三個實例來闡述群集檢測方法之應用。這三個實例為 (1) 本文作者與種苗改良繁殖場合作的實例一從成本效益考慮馬鈴薯種薯檢定病害計畫中，探討群集檢測方法之最佳群集大小；(2) 本文作者與農業試驗所合作的實例一從群集檢測結果估算種子傳毒率 (seed

transmission rate)，簡稱種傳率；(3) 本文作者在 2009 年接受動植物防疫檢疫局諮詢實例一在檢疫樣本中偵測線蟲之感染率；而其中第三個實例雖然實際上並非使用群集檢測來操作，但它是藉用群集檢測之概念，利用本文方程式 (1) 來完成計算，因此我們將一併納入討論，最後並探討群集檢測未來在植物病理學領域之可能發展方向。

群集檢測之原理及方法

群集檢測資料型式為二元隨機變數 (binary)，即每次檢測結果為陽性或陰性 (例如 ELISA)；在群集檢測結果中，如為陰性，表示此檢測群集每個單位皆為陰性；若為陽性，表示在這檢測群集中至少有一單位為陽性，但不知何者為陽性且有多少單位為陽性。本文使用與 Hughes 與 Gottwald 之報告⁽⁶⁾ 相同符號，假若單株染病之機率為 p_{low} ，當使用 n 株為一群集當成檢測對象時，在這群集中至少有一株罹病之機率為 p_{high} ，它們的關係如下：

$$p_{high} = 1 - (1 - p_{low})^n \quad (1)$$

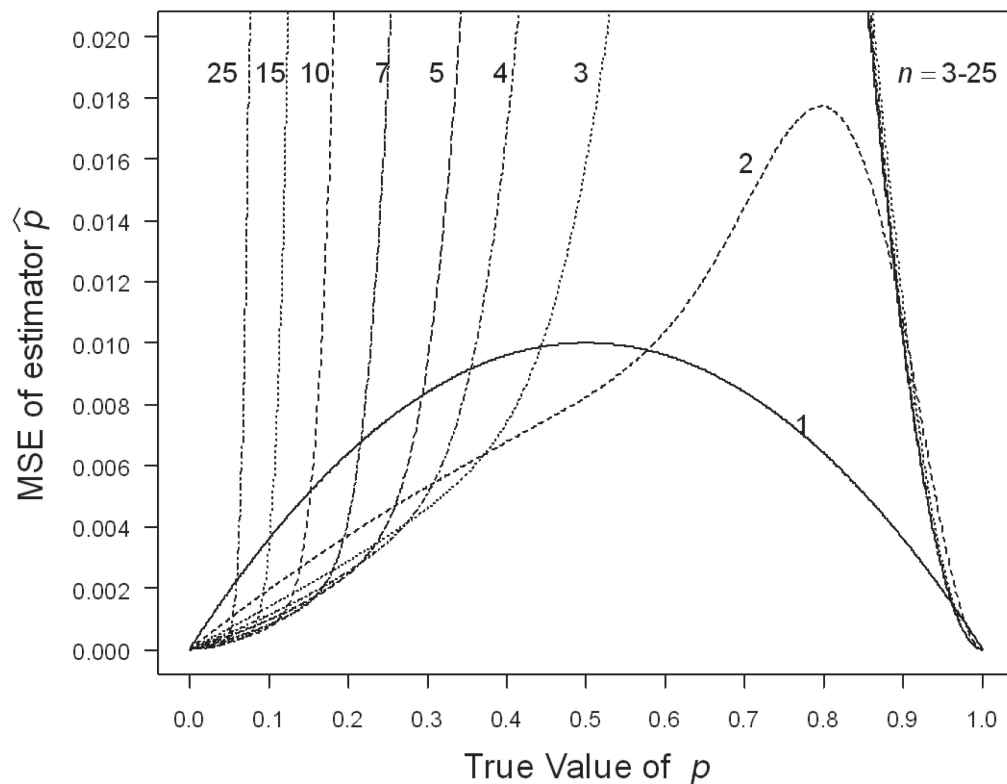
上式需有兩個假設前提，第一為單株間罹病機率應為獨立，即不會因距離的遠近影響罹病機率；第二為所

使用之檢測法 (例如 ELISA) 被視為無任何檢測誤差。

通常群集檢測有兩個目的，首先是估計在族群中檢測出陽性之機率 (即 p_{low})，它能提供植物保護工作者在田間管理決策上之依據；一般田間調查我們首先所獲得的資訊僅有 p_{high} ，即以群集為單位之測量值，可使用方程式 (1) 推得 p_{low} 數值，此類型問題稱之估計 (estimation)，在文後所舉例子以群集檢測估算種子傳毒率即為此種類型。群集檢測另一目的，即為分類 (classification)，在文後所提台灣種植之馬鈴薯病毒感染率之檢測即屬於此類，因它的目的是生產無毒種苗，故必須將有罹病之種苗挑出，當群集為檢測單位被測出病毒，必須再針對這一檢測為陽性的群集重新測試 (retest)，如果重新測試是採用逐一檢測，則稱為 Dorfman screening。近來有許多發表的報告，闡述當重新檢測施行時，逐一檢測並非最佳，因而提出許多新的方法，但本文著重在群集檢測概念之介紹與推廣，且探討這類不同重新檢測方法之報告大多發表於醫學領域⁽¹⁰⁾，故於此不擬著墨太多細節。

群集檢測分類最重要的問題首推最佳的群集大小 (n) 為何？一般 n 與 p_{low} 是有相關的， p_{low} 越小致使 n 值

越大⁽²⁰⁾，但 p_{low} 是族群未知的參數、亦是我們想估計探究之值，無法獲知其真正的大小。對此 Swallow 建議使用 p_{low} 的上限⁽²⁰⁾，即事先猜測 p_{low} 最大值為何，而在這篇報告中所使用的判斷準則是均方誤差 (mean squared error, MSE)，MSE 等於偏差 (bias) 的平方再加上變異數 (variance)，因此它可提供感染率估計值準確度 (accuracy) 與精確度 (precision) 之標準。舉例來說，圖一代表在 25 個群集 ($N = 25$) 情況下，每群集有 n ($= 1, 2, 3, 4, 5, 7, 10, 15, 25$) 棵植株時之真實感染率 (p) 與其估計式 (estimator) 的 MSE 相對應關係圖，可發現在 $n \geq 2$ 時，有些 p 值對應之 MSE 皆小於 $n = 1$ (逐一檢測) 之 MSE；在 $n = 2$ 的情況， p 值小於 0.6 時， $n = 2$ 之 MSE 皆小於 $n = 1$ 之 MSE，這點與我們的直覺認知不盡相同，通常我們會認為如用群集檢測，應可預知節省成本與時間，那可能會讓感染率估計值之準確度與精確度下降，但從圖一告訴我們，當 p 值小時，這情況將不會發生。但是， p 值需要多小，才能安心使用群集檢測？Venette、Moon 與 Hutchison 推薦 p 值小於 0.2 可能是較佳適用範圍⁽²³⁾。Swallow 指出群集數目 (N)、真實感染率 (p) 與群集大小 (n) 之間的關係^(20, 21)，當 N



圖一、每群集有 n ($= 1, 2, 3, 4, 5, 7, 10, 15, 25$) 棵植株且共有 25 個群集數目情況下，單一植株感染率相對應於其估計量均方誤差之關係圖。

Fig. 1. Mean squared error of the maximum likelihood estimator (\hat{p}) of the infection rate versus the true value of p for group sizes employing $n = 1$ to 25 with N (the number of groups) = 25.

值變大或 p 值變小，其 n 值變大，但 n 值不能無限擴大，在田間試驗中， $n = 50$ 應是最大上限。這其中除了田間調查方便外，最重要是考慮到檢測方法之敏感度 (sensitivity)，敏感度定義為當所檢測對象真正是陽性反應時，檢測結果為陽性反應之機率，當 n 太大時，可能有稀釋 (dilution) 效應產生，導致敏感度下降，影響檢測品質；Swallow 同時提及，若事先高估了 p 值，其 MSE 提高有限，但若低估 p 值，MSE 會急遽增大，導致所估數值的品質堪慮，因此建議儘量高估 p 值^(20, 21)。假使在實際情形下，事先沒有任何 p 值的資訊，則可使用逐次取樣 (sequential sampling)。澳洲學者在研究康乃馨病毒病害之檢測，曾提出使用三段逐次取樣來處理此一問題^(4, 18)，其意義為先用較大的群集大小 (例如 $n = 25$)，如果檢測結果全為陽性，再縮小群集大小 (例如 $n = 5$) 來做檢測；若不全為陽性，繼續以此群集大小操作。此外群集檢測分類問題，最佳的群集大小也可用成本來衡量，在接下來的實際應用案例部分，將會以本文作者與種苗改良繁殖場合作之報告⁽²⁾ 為例，詳述此過程。

實際應用案例

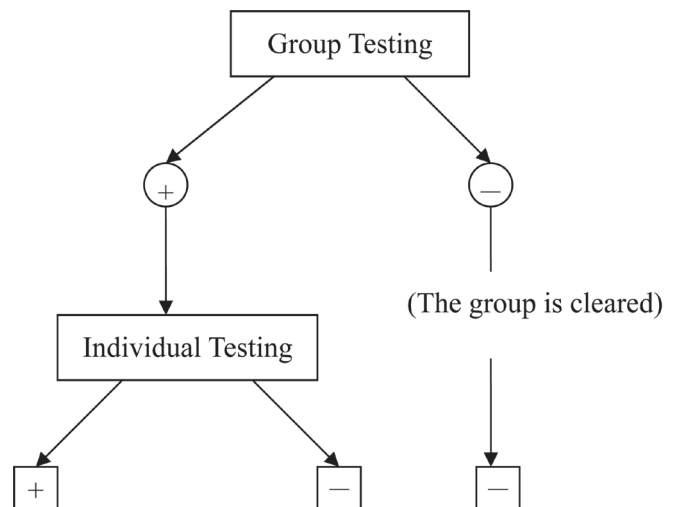
馬鈴薯病毒病感染率之群集檢測

台灣馬鈴薯的產地主要集中於台中、雲林與嘉義等三縣，其中以雲林縣斗南地區栽培面積最大，其次為台中縣豐原與嘉義縣溪口地區。由於台灣地處熱帶與亞熱帶地區，高溫多濕，馬鈴薯於栽培期間易遭受許多病蟲害的危害，其中又以病毒病較嚴重，因其可藉由無性繁殖體傳播，進而快速且廣泛的影響整個馬鈴薯之產業發展，故受到產業界、官方機構與研究機關極度重視。馬鈴薯之病毒病害種類多達 25 種，而大多數生產馬鈴薯的國家所遭遇的主要病毒病害，以馬鈴薯 A 病毒 (Potato virus A, PVA)、馬鈴薯 S 病毒 (Potato virus S, PVS)、馬鈴薯 X 病毒 (Potato virus X, PVX)、馬鈴薯 Y 病毒 (Potato virus Y, PVY) 及馬鈴薯捲葉病毒 (Potato leaf roll virus, PLRV) 為主⁽¹²⁾，因此台灣馬鈴薯健康種薯生產之標的檢測對象亦是針對此五類病毒進行。行政院農業委員會為獲得優良的健康馬鈴薯種薯，持續供應農民生產，特委由種苗改良繁殖場設立健康馬鈴薯種薯生產三級制，即從組織培養、基本種至原原種階段均由種苗改良繁殖場生產，至於原種與採種兩階段則分別由各地區農會或合作社等單位進行量產的工作。

現今在種苗改良繁殖場是採用 20 株馬鈴薯植株當

作一個群集 (group) 檢測單位，但往往造成成本過高，因此是否可讓每個群集有最適數目的植株數，且不影響檢測之精準度，將是主要探討的重點。本研究共在 2004、2005、2006、2007 與 2008 年 3 月 5 日進行馬鈴薯田間病毒之檢測，在檢測過程中針對 20 株之每株馬鈴薯隨機抽取一葉片，將此 20 個葉片混合研磨，進行直接酵素聯結抗體免疫法 (direct-ELISA)，檢測上述之五種病毒是否為陽性反應。因本研究旨在探討每個群集所包含植株之數量對其最終精準度之影響，故將其檢測結果視為健康或罹病，並不特別針對每類病毒加以探討。

田間與實驗室之操作細節請詳見本文作者發表於 2008 年之報告⁽¹⁾，此群集檢測計畫流程如圖二，如同上述介紹的 Dorfman screening，該報告中的最佳樣本數判斷準則為 MSE，這是從所估感染率之準確度與精確度觀點而言；但如從成本效益來考量可能更符合農民之需求，因此改從成本效益考量來探究最佳群集大小之計算。根據這套發展的計算流程⁽²⁾，只要給予所欲栽種的株數及感染率可能的上限，即可算出在多少群集大小情況下能使成本最少，它可適用於任何作物與地區，只要檢驗之結果為陽性或陰性 (即二分法)；在此研究中我們使用過去 5 年資料，感染率最大值為



圖二、兩階段的群集檢測方法

⊕：指出所檢測的群集為陽性。

⊖：指出所檢測的群集為陰性。

⊕：指出所檢測的個體單位為陽性。

⊖：指出所檢測的個體單位為陰性。

Fig. 2. Two-stage group testing design

⊕：Indicates positive result on the group scale.

⊖：Indicates negative result on the group scale.

⊕：Indicates positive result on the individual scale.

⊖：Indicates negative result on the individual scale.

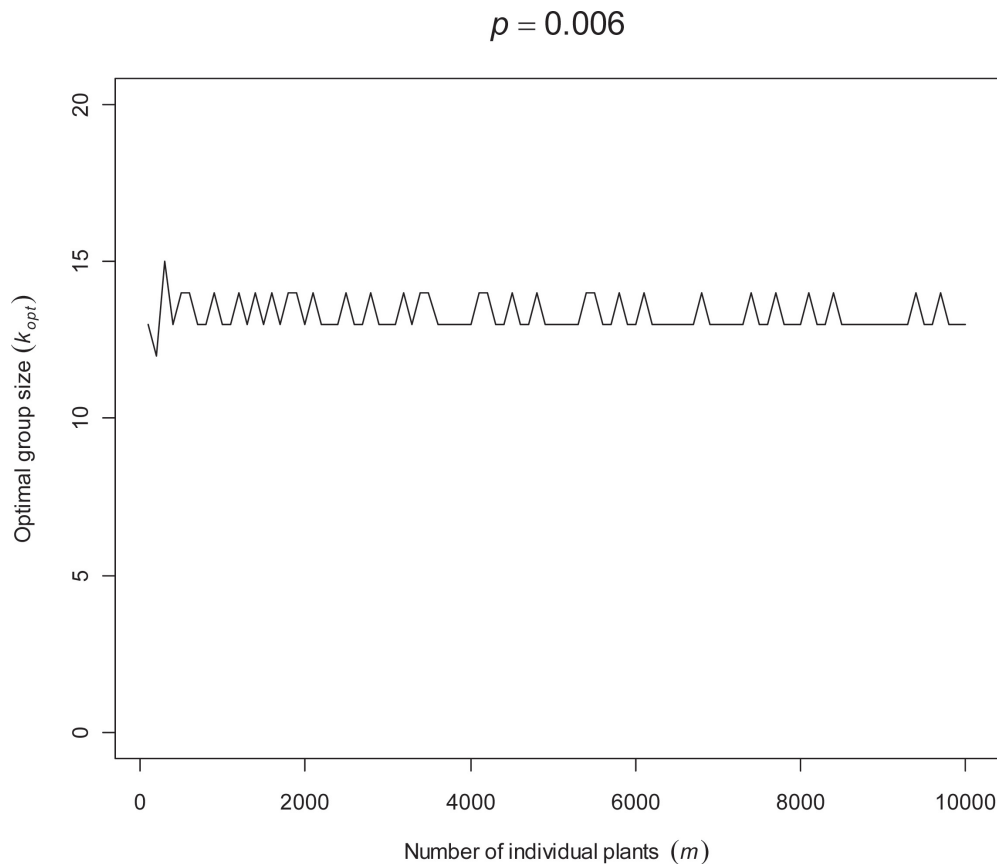
0.006；最佳群集大小與所栽種植株數關係如圖三，結果關係差異不大。最佳群集大小與感染率之關係如圖四，當感染率越大時，最佳樣本數就越小。本研究因在網室栽培，故感染率極小甚至有時為 0，但卻不能用太小的感染率，否則檢測敏感度可能下降且田間試驗群集大小甚少採用超過 50⁽²⁰⁾，因此採用歷年感染率最大值 0.006 是合理的，結果算出最佳群集大小為 13。表一顯示群集大小使用 20 與 13 之區別，其結果差距相當有限（僅 1.1 倍）；但與逐一檢測比較則有相當優勢，逐一檢測成本為其 6.6 倍（每個群集單位檢測成本大約 NT\$190 元），由此可揭示出族群檢測再分類之應用，可在最低成本下挑出有病毒病害的種薯。

以群集檢測結果估算種子傳毒率

建立種子健康檢查的標準流程，除了檢測技術的開拓外，對於病毒種傳率之最高容許度 (tolerance level) 必須充分瞭解，才能與國際的植物健康驗證 (phytosanitary certification) 制度接軌。而種傳率估計值是執行種傳忍受極限之重要依據，因此此部分主要是

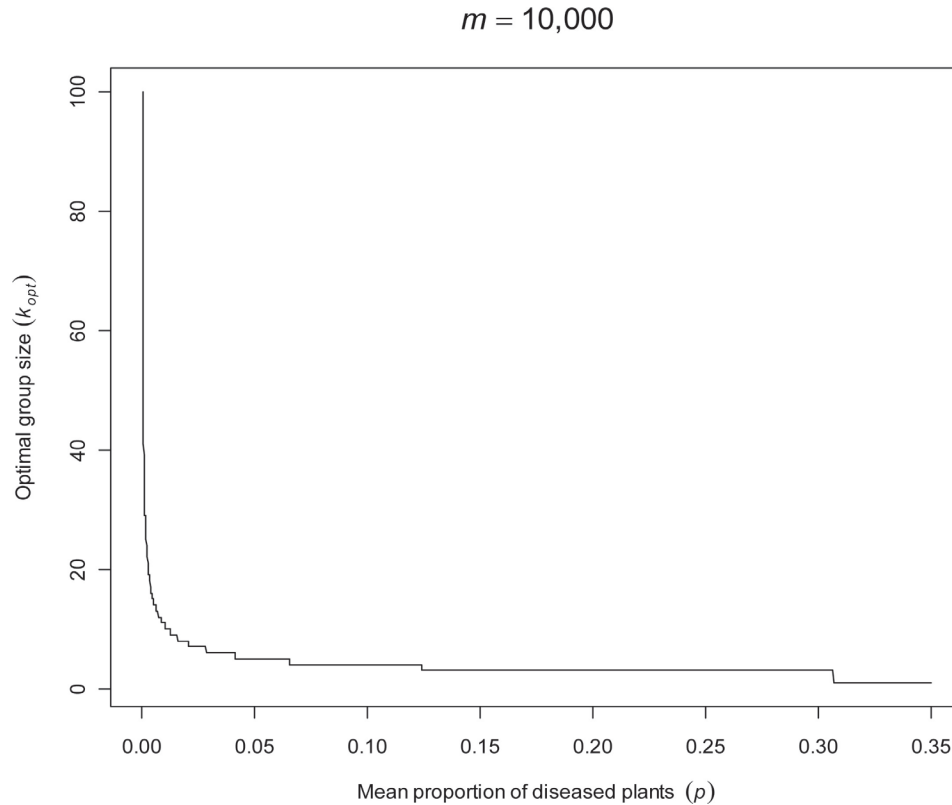
探討並發展一種統計分析方法來估算群集檢測方法的種傳率。

在計算種傳率時，先將待測之供試樣本種子數依 n 值加以區隔成 N 個群（即在每個群內有 n 粒種子），每一群內所有種子樣本集合共同萃取出單一樣本進行檢測。利用英國園藝研究機構 (Horticulture Research International) 在 1995 年所開發之種子試驗分析程式 (seed test analysis program, *STpro*)⁽¹⁷⁾，可估算出種傳率 (θ) 之點估計值、第 i 個樣本之陽性反應配適值 (fitted value)、適合度檢定 (goodness of fit)，並以最大概度估計法 (maximum likelihood estimation) 之大樣本變異數來建構種傳率的 95% 信賴區間。此處與先前群集檢測之不同，在於群集大小 (n) 並非全然相同。例如在扁蒲種子檢測中，想探討胡瓜綠斑嵌紋病毒 (*cucumber green mottle mosaic virus*, CGMMV) 之傳毒率， n 值分布從 1, 2, 5, 10 至 100， n 值不同會讓群集檢測複雜化，之前亦有文獻探討^(4,5) 並著重於感染率信賴區間之估算；本文作者亦針對此情形討論過度離勢 (overdispersion) 情形時之感染率信賴區間估算⁽¹¹⁾，並將此寫成容易使用



圖三、當感染率為 0.006 時，最佳群集大小與所栽種植株數目之關係。

Fig. 3. Relationship between optimal group size (k_{opt}) and the number of individual plants (m) for the mean proportion of diseased plants (p) = 0.006.



圖四、當所栽種植株數目為 10,000 株時，最佳群集大小與感染率之關係。

Fig. 4. Relationship between optimal group size (k_{opt}) and the mean proportion of diseased plants (p) for the number of individual plants (m) = 10,000. The values of p chosen are 0.0001, 0.0002, ..., 0.3500.

表一、當使用群集檢測在馬鈴薯種苗驗證計畫時，群集大小 13 株與 20 株平均成本之比較，13 株為使用自行撰寫的程式 (“group-size”) 算出，20 株為種苗改良繁殖場在 2004-2008 年取樣時通用的標準

Table 1. Comparison of average cost (NT \$) per plant between $n = 13$ and $n = 20$. The former group size is obtained from the code “group-size” and the latter was used by the Seed Improvement and Propagation Station in Taiwan during the period of 2004 - 2008

m^1	p^2	Average cost per plant		cost (20) / cost (13) ³
		$n = 13$	$n = 20$	
2,000	0.006	28.91634	31.04602	1.073650
5,000	0.006	28.91946	31.04602	1.073534
10,000	0.006	28.92477	31.04602	1.073337

¹ the number of individual plants

² the mean proportion of diseased plants

³ cost (n) represents the required cost when the group size equals n .

之程式供下載使用，它可解決配適不好之情形，且可提供感染率較佳之信賴區間估計值。圖五是以與農業試驗所合作扁蒲種子 CGMMV 傳毒率的例子，而使用類最大概度估計法 (quasi-likelihood estimate) 來解決過度離勢現象，因為偏斜校正計分法 (skewed-corrected score) 經模擬 (simulation) 檢驗最為穩定⁽¹⁾，故使用此方法來建構信賴區間。

在檢疫樣本中偵測線蟲之感染率

2009 年澳洲欲對台灣輸出紅蘿蔔，針對穿孔線蟲 (burrowing nematode) 檢疫問題提出抽樣的方法，在這份文件中提及在每單位面積中增加採樣的數目並無法增加偵測到線蟲的機率，它引用 McSorley 與 Littell 的文章，在 1993 年報告所使用的方程式即為本文方程式 (1)，先決定預估感染的機率 (即 p_{low})，再加上所需抽樣

```

> k=c(1,1,1,2,2,2,5,5,5,10,10,10,100,100,100)
> N=c(10,10,10,10,10,10,10,10,10,10,10,10,5,5,5)
> r=c(0,1,0,1,2,0,0,1,1,2,4,0,0,0,0)
> seed.health(k,N,r,0.05)
[1] "Welcome to seed.health"
$Output
      No. seeds in sample No. samples tested No. positive Fitted value
[1,]                1                10                0      0.0598
[2,]                1                10                1      0.0598
[3,]                1                10                0      0.0598
[4,]                2                10                1      0.1192
[5,]                2                10                2      0.1192
[6,]                2                10                0      0.1192
[7,]                5                10                0      0.2954
[8,]                5                10                1      0.2954
[9,]                5                10                1      0.2954
[10,]               10                10                2      0.5820
[11,]               10                10                4      0.5820
[12,]               10                10                0      0.5820
[13,]              100                 5                0      2.2550
[14,]              100                 5                0      2.2550
[15,]              100                 5                0      2.2550

$`P-value of goodness-of-fit chi-squared statistic`
[1] 0.9201

$`The estimated dispersion parameter`
[1] 6.6724

$`The estimated proportion of infected seeds`
[1] 0.006

$`The nominal level(%)`
[1] 95

$`The lower confidence limit`
[1] 0.0011

$`The upper confidence limit`
[1] 0.0172

```

圖五、利用所發展群集測試方法估算種子傳毒率的程式 (“*seed.health*”)，並且使用扁蒲種子 (1877T-2B) 檢測胡瓜綠斑嵌紋病毒 (CGMMV) 為例，此處 k 即為文章中 n 值代表每一群內種子數， N 代表所測試群集個體數， r 代表測試出為陽性的群集數。

Fig. 5. Example output from the “*seed.health*” program. Data are from tests of the seed lot (1877T-2B) for the determination of seed-transmission rates of *Cucumber green mottle mosaic virus* (CGMMV) in bottle gourds. Here, k , N , and r represent the group size, the number of groups, and the number of positive groups, respectively.

樣本數 (即為 n 值)，可求得偵測出線蟲感染率之百分比⁽¹⁶⁾。例如在一塊田區有 5% 線蟲的感染率下，若抽取 59 個樣本可偵測到感染率為 95%，若抽取 90 個樣本則偵測到感染率為 99%，即使樣本數超過 100，仍幾乎無法提高線蟲偵測的機率。因此澳洲當局用此論點來說明，增加太多取樣點其實對提高線蟲偵測的機率並沒有助益。雖然此類問題實質上並非群集檢測，但概念

上卻是使用群集檢測的方法。再者，1993 年這篇報告中的假設似過於嚴格，即假設線蟲分布是隨機且取樣方法亦是隨機取樣，若線蟲在田間分布非隨機，恐不適用於該報告之應用範圍；另外，文中陳述在感染的單位中需 100% 能被偵測出線蟲存在，這在實際情況通常不成立，例如偵測土壤中的線蟲就常無法達到 100% 的偵測效率。

群集檢測在植物病理學研究之未來發展

群集檢測方法在植物病理學門相關研究的未來發展，特此介紹如下：首先，群集檢測假設每個個體 (如植株) 彼此是獨立地罹病，但在真實情況此假設可能不成立，例如 Hughes 與 Gottwald 指出柑橘南美立枯病的媒介昆蟲為大桔蚜 (*Toxoptera citricida*)⁽⁷⁾，其為群聚性而非彼此獨立，故本文方程式 (1) 的等號右邊將須作改變，而使用有效樣本數 (effective sample size) 可解決此一問題，即若原來每群集的樣本數為 20，但因群聚性，可能導致僅有 15 個獨立樣本；該縮減變化是由群聚參數 (aggregated parameter) 在控制，當群聚參數越大，有效樣本數會變得更小^(13, 14)。此外，近年來植物病理學界有些報告提出階層取樣 (hierarchical sampling) 的概念^(8, 9, 15, 19, 22)，其實它是群集檢測的概念延伸；傳統群集檢測視單一個體彼此之間為獨立 (例如田區每株植株罹病率)，故使用二項分布 (binomial distribution) 來量化田間罹病狀況；但若為階層取樣，單一個體彼此之間呈現非獨立現象，即群聚現象產生，必須以貝他二項分布 (beta-binomial distribution) 來取代二項分布描述田間植株罹病情況。

第二，群集檢測的假設是檢測方法 (例如 ELISA) 沒有任何檢測誤差，亦即沒有錯誤分類 (misclassification)，但實際情形卻是不可能的，因此如何估算其誤差，且在有誤差情況下估計值變化情形為何，正是本文作者目前進行中的研究。此類研究在醫學與動物流行病學有頗多的討論⁽¹⁰⁾，其實不管對象是動物或植物，在實驗室中檢測方法原理是雷同的，在不同學門間統計方法的整合將可解決此問題。

結 論

群集檢測常使用在植物病理學的研究，在台灣亦有不少運用實例。本文最重要目的為整理與說明該法之原理及實用性，並引用本文作者最近三年來與台灣植物病理專家合作運用該法的經驗，希望能達到拋磚引玉之效果，為植物保護盡一份心力。

群集檢測可證明單株檢測的精確度並非永遠最高，當單株感染率甚小時尤其明顯，此方法並不侷限於本文所描述的應用實例。對於各類病害的調查，當其感染率甚小且檢測成本所費不貲時，在取樣上群集檢測當可發揮其最大效能。使用群集檢測時，切勿使用過大的群集大小，否則將導致所估的感染率產生極大的誤差，且出現偽陰性 (false-negative) 的機率也會升高，最終影響檢測的準確性。

謝 辭

本文承蒙兩位審查者提供意見及斧正，謹此致謝。

引用文獻 (LITERATURE CITED)

1. Chiang, K. S., Chung, W. C., Lin, S. H., Lai, S. H., and Wang, S. F. 2008. Group-testing design for the infection rates of potatoes viruses. *Plant Pathol. Bull.* 17: 321-326. (in Chinese).
2. Chiang, K. S., Lai, S. H., Chung, W. C., Lin, S. H., and Yang, T. C. 2010. Cost considerations in choosing group size for group testing in the seed potato certification program. *Am. J. Potato Res.* (In press)
3. Dorfman, R. 1943. The detection of defective members of large populations. *Ann. Math. Stat.* 14: 436-440.
4. Hepworth, G. 1996. Exact confidence intervals for proportions estimated by group testing. *Biometrics* 52: 1134-1146.
5. Hepworth, G. 2005. Confidence intervals for proportions estimated by group testing with groups of unequal size. *J. Agric. Biol. Envir. S.* 10: 478-497.
6. Hughes, G., and Gottwald, T. R. 1998. Survey methods for assessment of citrus tristeza virus incidence. *Phytopathology* 88: 715-723.
7. Hughes, G., and Gottwald, T. R. 1999. Survey methods for assessment of citrus triteza virus incidence when *Toxoptera citricida* is the predominant vector. *Phytopathology* 89: 487-494.
8. Hughes, G., Gottwald, T. R., and Levy, L. 2002. The use of hierarchical sampling in the surveillance program for *Plum pox virus* incidence in the United States. *Plant Dis.* 86: 259-263.
9. Hughes, G., McRoberts, N., Madden, L. V., and Gottwald, T. R. 1997. Relationships between disease incidence at two levels in a spatial hierarchy. *Phytopathology* 87: 542-550.
10. Kim, H. Y., Hudgens, M. G., Dreyfuss, J. M., Westreich, D. J., and Pilcher, C. D. 2007. Comparison of group testing algorithms for case identification in the presence of test error. *Biometrics* 63: 1152-1163.
11. Liu, S. C., Chiang, K. S., Lin, C. H., and Deng, T. C. 2010. Confidence interval procedures for proportions estimated by group testing with groups of unequal size adjusted for overdispersion. *J. Appl. Stat.* (In press)
12. Loebenstein, G. 2007. Potato virus diseases, their diagnosis and preparation of virus-tested seed potatoes by rapid propagation. *Acta Hort. (ISHS)* 729: 437-440.
13. Madden, L. V., and Hughes, G. 1999. An effective sample size for predicting plant disease incidence in a spatial hierarchy. *Phytopathology* 89: 770-781.

14. Madden, L. V., and Hughes, G. 1999. Sampling for plant disease incidence. *Phytopathology* 89: 1088-1103.
15. Madden, L. V., Hughes, G., and van den Bosch, F. 2007. Estimating plant disease sampling. Pages 279-318 *in: The Study of Plant Disease Epidemics*. American Phytopathological Society, St. Paul, MN, 432 pp.
16. McSorley, R., and Littell, R. C. 1993. Probability of detecting nematode infestations in quarantine samples. *Nematropica* 23: 177-181.
17. Ridout, M. S., and Roberts, S. J. 1995. Instructions and notes on *STpro* program. Horticulture Research International, UK. 8 pp. (Website: www.planthealth.co.uk/downloads).
18. Rodoni, B. C., Hepworth, G., Richardson, C., and Moran, J. R. 1994. The use of a sequential batch testing procedure and ELISA to determine the incidence of five viruses in Victorian cut-flower Sim carnations. *Aust. J. Agric. Res.* 45: 223-230.
19. Shah, D. A., Dillard, H. R., and Nault, B. A. 2005. Sampling for the incidence of aphid-transmitted viruses in snap bean. *Phytopathology* 95: 1405-1411.
20. Swallow, W. H. 1985. Group testing for estimating infection rates and probabilities of disease transmission. *Phytopathology* 75: 882-889.
21. Swallow, W. H. 1987. Relative mean squared error and cost considerations in choosing group size for group testing to estimate infection rates and probabilities of disease transmission. *Phytopathology* 77: 1376-1381.
22. Turechek, W. W., and Madden, L. V. 2003. A generalized linear modeling approach for characterizing disease incidence in a spatial hierarchy. *Phytopathology* 93: 458-466.
23. Venette, R. C., Moon, R. D., Hutchison, W. D., 2002. Strategies and statistics of sampling for rate individuals. *Annu. Rev. Entomol.* 47: 143-174.



Dr. Chiang, K. S.

Dr. Chiang is currently an assistant professor in the Department of Agronomy, at National Chung Hsing University (NCHU), Taichung, Taiwan. In 2000, he earned a Ph.D. degree from the University of Pittsburgh in the USA with a major in Biostatistics and a minor in Epidemiology. His major expertise is in statistical methodologies and applications in agriculture and natural resources with special emphasis on the statistical methods of plant protection. In recent years, he has focused on the use of statistical methods in plant epidemiology. Currently, he has been cooperating with some plant pathologists in Taiwan for developing and applying group testing procedures with positive results. Also, he has actively participated in several international conferences to communicate with experts abroad to widen his range of knowledge.



Mr. Lai, S. H.

Mr. Lai received his B. S. and M. S. degrees in 2004 and 2007, respectively from the Department of Agronomy, at National Chung Hsing University (NCHU), Taichung, Taiwan. He is currently a Ph. D. student under the direction of Dr. Chiang working on population fluctuations and forecasting population densities of brown planthoppers and rice leaffolders.

ABSTRACT

Chiang, K. S.^{1,2}, and Lai, S. H.¹. 2010. Application of group testing procedure in plant pathology. *Plant Pathol. Bull.* 19: 117-126. (¹ Department of Agronomy, National Chung Hsing University, Taichung, Taiwan; ² Corresponding author, E-mail: kucst@dragon.nchu.edu.tw; Fax: +886-4-2285-6497)

Group testing is a method of pooling a number of units together and performing a single test on the resulting group. The procedure is particularly useful when the number of positive units is expected to be low and obtaining test material is cheap, but testing itself is expensive. The method was introduced into the field of plant pathology mainly in a pioneering study by Swallow⁽²⁰⁾. Since then, many studies regarding group testing have been published in leading journals of plant pathology. The results of these studies indicate that group testing offers dramatic improvements in precision and accuracy of the sampling information, with no undue increases in costs or labor. Therefore, it is imperative for plant pathologists in Taiwan to become familiar with group testing. The objectives of this review article are first, to explain group testing methods and procedures. Secondly, three examples are used to illustrate the application of group testing to real plant pathology conditions in Taiwan. Finally, the further development of group testing in plant pathology is discussed.

Key words: group testing, estimation, classification, mean squared error, cost