

國立中興大學農藝學系生物統計研究室

Biometrics Laboratory

Department of Agronomy, National Chung-Hsing University



指導老師：郭寶錚教授 (Bo-Jein Kuo, Ph.D.)

學歷：美國 Virginia Commonwealth Univ. 生物統計系博士

美國 Iowa State Univ. 統計系碩士

國立中興大學糧食作物研究所碩士

國立中興大學農藝系學士

研究領域：生物統計、試驗設計、NIR 光譜資料處理、淨最小平方方法之研究、生物晶片之統計方法、模擬基改作物花粉飄散及量測不確定度的評估

研究室簡介：

生物統計學是研究生命科學、醫學及農學研究的重要工具。郭寶錚教授開設課程有大學部的「生物統計學」、「回歸與相關」及「試驗設計學與實習」課程；研究所的「回歸分析」課程，這些課程能作為研究生規劃試驗及撰寫碩博士論文之工具。本研究室以培育生物資訊統計等跨領域人才為宗旨，研究領域包括生物統計、試驗設計、NIR 光譜資料處理、淨最小平方方法之研究、生物晶片之統計方法、模擬基改作物花粉飄散及量測不確定度的評估，結合理論與應用兩大面向，以期學生不只是死讀書，更能學以致用。

開設課程：

1. 回歸與相關 (Regression and Correlation)
2. 生物統計學 (Biometrics)
3. 回歸分析 (Regression Analysis)
4. 試驗設計學與實習 (Experimental Design and Practice)
5. 生物統計專題討論 (Seminar in Biometrics)
6. 統計方法與試驗設計專題討論 (Seminar on Statistical Method and Experimental Design)
7. 高等統計方法與試驗設計專題討論 (Seminar on Advanced Statistical Method and Experimental Design)
8. 農藝學研究法 (Research Methods for Agronomy Science)

近 5 年主持計畫：

104~107 年，國科會，模擬基改油菜花粉飄散之研究

104 年，農糧署，偵測大小油菜間基因流動形態標誌之建立

104 年~107 年，種苗場研究計畫

近 5 年共同主持計畫：

106 年，農委會，精準/無農藥殘留茶園栽培管理技術開發

107~110 年，科技部，利用 AI 智慧感知技術建構溫室番茄栽培管理專家系統

107~108 年，科技部，結合 UAV 監測之智慧農業栽培支援系統

近 5 年發表論文：

1. 王柏盛、蘇圓智、郭寶錚 2015 利用不同統計方法建立花飄散模式 中華民國雜草學會會刊 36(1): 33-62
2. 余昇驊、郭寶錚 2016 模擬在臺灣地景等級下基改玉米花粉媒介基因流布及共存措施 作物、環境與生物資訊 13: 67-79
3. 洪鴻、林子凱、杜元凱、郭寶錚 2016 模擬 GM 大油菜與小油菜之雜交後代的鑑別 作物、環境與生物資訊 13: 53-66
4. 連柏雁、余昇驊、郭寶錚 2016 以高斯煙羽—gamma 法模擬在臺灣地景等級下基改玉米集中種植之共存措施 作物、環境與生物資訊 13: 105-115
5. 洪鴻、林子凱、杜元凱、郭寶錚 2016 利用 SRAP 分子標誌進行 *Brassica rapa* 與 *Brassica napus* 及其雜交後代的鑑別 作物、環境與生物資訊 13:208-222
6. 連柏雁、郭寶錚 2016 使用不同統計模式推估基改與非基改玉米共存之隔離距離 作物、環境與生物資訊 13: 25-38
7. 楊芷翎、王柏盛、蘇圓智、郭寶錚 2017 短距離下蟲媒基改油菜花粉飄散及相關影響因素介紹 作物、環境與生物資訊 14: 5-16
8. 陳思婷、許鑄云、郭育奴、郭寶錚 2018 多次調查減緩種子試驗中不足變異之評估 作物、環境與生物資訊 15:159-168
9. Laffont, J. L., B. Hong, B. J. Kuo & K. M. Remund. 2019. Exact theoretical distributions around the replicate results of a germination test *Seed Science Research*, 29, 64-72.

近 5 年國內外研討會：

1. 洪鴻、郭寶錚、古新梅。2015。利用 SRAP 分子標誌進行 *Brassica rapa* 與 *Brassica napus* 及其雜交後代的鑑別。台灣農藝學會 104 年度年會作物科學講座暨研究成果發表會，屏東。
2. 廖莉婷、林汶鑫、郭寶錚。2015。Evaluation of the uninformative variable elimination (UVE) method in variable selection for NIR data。台灣農藝學會 104 年度年會作物科學講座暨研究成果發表會，屏東。
3. 蘇圓智、洪鴻、林子凱、郭寶錚。2015。Effect of wind direction and distance on gene flow between *Brassica napus* L. and *Brassica campestris* L. in Taiwan。台灣農藝學會 104 年度年會作物科學講座暨研究成果發表會，

- 屏東。
4. Yi-ting Zhang, Han-Tsu Shen and Bo-jein Kuo. 2016. Comparison of Different Top-Down Strategies for Measurement Uncertainty Estimation in Quantification of Genetically Modified Organisms. ISTA Congress 2016, Estonia.
 5. 李承彬、郭寶錚。2018。以柯西加權模式模擬玉米花粉調控之基因流布。台灣農藝學會 107 年作物科學講座暨研究成果發表會，嘉義。
 6. 陳思婷、郭寶錚。2018。種子發芽試驗中不足變異之探討。台灣農藝學會 107 年作物科學講座暨研究成果發表會，嘉義。
 7. 楊芷翎、郭寶錚。2018。利用零膨脹負二項模式擬基因造改大油菜對小油菜之花粉飄散距離 — 以台中霧峰地區為例。台灣農藝學會 107 年作物科學講座暨研究成果發表會，嘉義。
 8. 廖鈺婷、郭寶錚。2018。以準確度曲線 IQC 程序評估基因改造作物定量分析。台灣農藝學會 107 年作物科學講座暨研究成果發表會，嘉義。
 9. 蔡政煌、郭寶錚。2018。在群組試驗中利用貝氏信賴區間檢定種子批之轉基因性狀比例。台灣農藝學會 107 年作物科學講座暨研究成果發表會，嘉義。
 10. Si-Ting Chen, Ciao-Yun Syu, Yu-Wen Kuo, Bo-Jein Kuo. 2019. The methods to prevent analysts unconsciously adjusting the result in the germination tests. ISTA Congress 2019, India.

Comparison of Different Top-Down Strategies for Measurement Uncertainty Estimation in Quantification of Genetically Modified Organisms



Yi-ting Zhang¹, Han-Tsu Shen² and Bo-jein Kuo¹
¹Department of Agronomy, National Chung Hsing University
²Taiwan Seed and Seedling Improvement and Propagation Station (TSIPS), C. O. A.



INTRODUCTION

Measurement uncertainty (MU) is defined as “non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand” [1]. MU may also be used as a reference for conformity assessments to demonstrate that whether the technical regulations or standard procedures are met by a certain product. The top-down strategy can directly estimate the uncertainty by calculating the reproducibility and repeatability. This study seeks to apply two top-down approaches, i.e. intermediate precision and repeatability standard deviation, to compare the performance in estimating the MU. The estimates of MU could serve as references for laboratory analysis personnel and stakeholders in decision-making.

MATERIALS & METHODS

1. Quantitative Analysis of GMOs

In this study, we used thirty samples for the proficiency test of the genetically modified (GM) events containing different concentrations of the powdered certified reference materials, including 0.5, 1, 10% for DP-305423 soybean (ERM-BF426) and 0.49, 0.98, 4.91% for NK603 maize (ERM-BF415), respectively. Ten DNA extracts of each sample were analyzed using the real-time PCR (StepOne Plus) purchased from Applied Biosystems (Thermo Fisher, USA). Each DNA extract was quantified in triplicate and expressed by the average of triplication. The quantification was repeated twice to achieve two independent estimates of the GM concentrations for the DP-305423 soybean (ERM-BF426) and NK603 maize (ERM-BF415), respectively.

2. Uncertainty Estimation

1) Outlier Detection

According to the recommendation of ISO 5725, prior to the estimation of uncertainty, Cochran's statistic should be used for the outlier testing [2].

2) Intermediate Precision Approach

Intermediate precision (RSD_p) may reflect the variation in a single laboratory; therefore, the methods proposed by Žel *et al.* (2007) [3] could be applied to the MU estimation. The RSD_p could be regarded as a combined standard uncertainty (u_c(y)). The expanded uncertainty (U) at the 95% confidence level was calculated as follows:

$$U = 2 \times RSD_p$$

3) Repeatability Standard Deviation Approach

The repeatability standard deviation method proposed by Ambrus (2004) [4] was estimated by duplicated analysis data. The repeatability standard deviation was defined as the uncertainty of the laboratory phase which could be further divided into “analysis” and “sample processing” parts, respectively. The uncertainty of analysis part was estimated by the recovery experiment. The uncertainty of sample processing was estimated by the law of propagation of uncertainty. In this study, the concentration of was certified reference materials regarded as the true concentration of the sample and could be used to determine the relative standard uncertainty of recovery.

3. Conformity Assessment

This study also utilized guard band and consulted 0.9% labeling threshold (according to European regulation) to evaluate the critical value of the decision, serving as a basis for conformity assessment [5]. The guard band was described as follows:

$$\text{guard band} = 0.9\% \times u_c(y) \times 1.65$$

Then, the critical value of the decision was derived in the following:

$$\text{threshold value} = 0.9\% \pm \text{guard band}$$

The higher limit defined as “high confidence of correct rejection” represented that stakeholder decision-makers would have higher possibility to determine the tested sample possessed GM content less than the maximum concentration of 0.9%. The lower limit defined as “high confidence of correct acceptance” and represented that stakeholder decision-makers would have higher possibility to determine the tested sample as possessing over 0.9% threshold concentrations of GM content.

RESULTS

1. Outlier Test

According to the recommendation of ISO 5725, the Cochran's statistical testing was used to determine the outliers in the dataset of DP-305423 soybean and NK603 maize, respectively. Seven outliers were deleted from two GM events. The deleted observations for two events tended to have a higher mean concentration (Fig. 1).

2. Intermediate Precision Approach

The relative expanded uncertainty (k = 2) values of GM soybean DP-305423 event and GM maize NK603 event estimated by the intermediate precision approach were 22.52% and 18.37%, respectively (Fig. 2).

3. Repeatability Standard Deviation Approach

The relative expanded uncertainty (k = 2) value of GM soybean DP-305423 event and GM maize NK603 event estimated by the repeatability standard deviation approach were 24.71% and 20.55%, respectively (Fig. 2). The primary source of variation was the sample preparation, accounting for 86.57% to 94.73% of the total variances (Fig. 3).

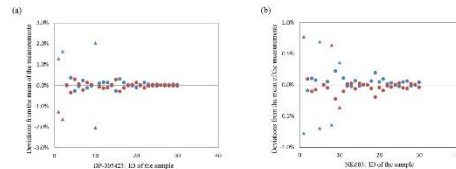


Fig. 1 Deviations from the mean of the measurements versus the ID of the sample. (a) DP-305423 (b) NK603. ● they are not outliers, ▲ they are outliers detected by Cochran's statistic; blue color represents result 1, red color represents result 2

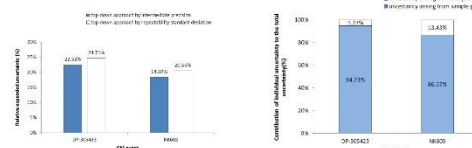


Fig. 2 Estimates of the relative expanded uncertainty by intermediate precision and repeatability standard deviation. Fig. 3 Estimates of the relative expanded uncertainty by repeatability standard deviation.

4. Conformity Assessment

Using the repeatability standard deviation approach, the critical value of the decision obtained from guard band indicated that two GM events with “high confidence of correct rejection” were 1.07% and 1.04%, respectively; events with “high confidence of correct acceptance” were 0.73% and 0.76%, respectively. Using the repeatability standard deviation approach, the critical value of the decision obtained from guard band indicated that two GM events with “high confidence of correct rejection” were 1.08% and 1.05%; events with “high confidence of correct acceptance” were 0.72% and 0.75% (Table 1).

Table 1 Critical value of the decision determined by the guard band (threshold value 0.9%).

GM event	Intermediate precision	
	High confidence of correct rejection	High confidence of correct acceptance
DP-305423	1.07%	0.73%
NK603	1.04%	0.76%
GM event	Repeatability standard deviation	
	High confidence of correct rejection	High confidence of correct acceptance
DP-305423	1.08%	0.72%
NK603	1.05%	0.75%

CONCLUSION

Two top-down estimation approaches resulted in consistent MU values of GM events. When the value of MU was too high, personnel could use the repeatability standard deviation approach to distinguish the main resource of variation in MU. Intermediate precision approach could be applied to other analyses with similar design of experiments. Lastly, these results can assist stakeholders or laboratory analysis personnel in decision-making.

References

- [1] International Organization for Standardization. General requirements for the competence of testing and calibration laboratories (ISO/IEC 17025:2005). Geneva, Switzerland: International Organization for Standardization; 2005.
- [2] International Organization for Standardization. Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method (ISO 5725-2:1994/Cor.1:2002). Geneva, Switzerland: International Organization for Standardization; 2002.
- [3] Žel J, Gruden K, Cankar K, Štebih D, Bletje A. Calculation of measurement uncertainty in quantitative analysis of genetically modified organisms using intermediate precision: a practical approach. J AOAC Int. 2007;90: 582-586.
- [4] Ambrus A. Reliability of measurements of pesticide residues in food. Accred. Qual. Assur. 2004;9:288-304.
- [5] Ellison SLR, Williams A. Use of uncertainty information in compliance assessment. St. Gallen, Switzerland: EURACHEM; 2007.

以準確度曲線與IQC程序 評估基因改造作物定量分析之量測不確定度



廖鈺婷 郭寶錚
國立中興大學農藝學系

A. 基因改造作物之定量分析與量測不確定度



- 近年來，隨著基因改造 (genetically modified, GM) 作物的普及，各國為管理GM作物及其衍生之問題，對於產品或食品因偶發而存在的GM成分含量上限值加以規範，然而GM作物之GM成分標示的門檻值各國不盡相同，其中歐盟為0.9%、韓國為3%、日本5%、台灣3%。當受測量物中GM含量超過規定門檻值則需標示含GM成分 (Monica *et al.*, 2014)，因此，定量分析中量測結果的精準度也就非常重要。
- 本研究以目前最常被應用於分析和偵測食品中是否存在GM成分的即時定量聚合酶鏈鎖反應 (real-time polymerase chain reaction, real-time PCR) 作為定量分析方法，並估計其量測不確定度值 (measurement uncertainty, MU)。
- MU是一個與量測結果相關的參數，用以呈現定量分析結果之變異性，並作為定量結果品質好壞的指標。本研究利用不同的MU評估方法，針對GM大豆品項DP-305423中三種不同濃度於定量分析後的資料進行評估，以做為相關領域分析與決策人員對於結果解釋的客觀依據。

B. 以準確度曲線評估GM作物定量分析之MU

- 為了評估一個分析方法的量測品質，以及其所造成之整體不確定度，除了利用內部品質控制程序作為資料分析的程序外，亦可利用方法確校 (method validation) 的策略來作為方法品質的評估。方法確校是一種藉由考慮系統誤差及隨機誤差來判定結果精確性的策略。
- 在定量分析程序中，對於分析方法的確校是否符合目的適性，可利用準確度曲線 (accuracy profile) 作為評估方法，也就是透過容忍區間來評估分析方法的結果值與真值間的絕對差值是否落在可接受的界限 (acceptability limit) 內，並以圖形來呈現結果的一種決策工具 (Saffaj *et al.*, 2013)。
- 本研究以β-ETI、β-CTI及MLS三種準確度曲線的評估方法，針對GM大豆DP-305423品項濃度為10%、1%、0.5%進行定量分析後MU的評估，結果顯示以β-ETI所估計的相對擴充不確定度值最大，並可發現在估算準確度曲線中的非中心卡方值與中心卡方值時，當後者大於前者時，以β-ETI估計的MU值會大於其餘兩種評估方法所估計的結果 (Table 1)。

Table 1 Estimates of the relative expanded uncertainty in GM soybean DP-305423 event by different accuracy profiles.

品項	濃度	非中心卡方值	中心卡方值	β-ETI 相對擴充不確定度 (%)	β-CTI 相對擴充不確定度 (%)	MLS程序 相對擴充不確定度 (%)
DP-305423	10%	1.0749	2.3372	21.3046	17.6359	17.5593
	1%	0.9936	9.1770	38.4499	28.3440	28.8336
	0.5%	0.9987	6.8040	33.5751	25.3390	25.5386

F. 參考文獻

- Monica, C. S. M., L. M. Gabriela, and G. Lorena. 2014. Adventitious presence of transgenic events in the maize supply chain in Peru: A case study. *Food Control* 41: 96-101.
- Saffaj, T., B. Ihssane, F. Jhilal, H. Bouchafra, S. Laslami, and S. Alaoui Sosse. 2013. An overall uncertainty approach for the validation of analytical separation methods. *Analyst* 138:4677-4691.

C. 以準確度曲線評估GM作物定量分析方法之方法確校

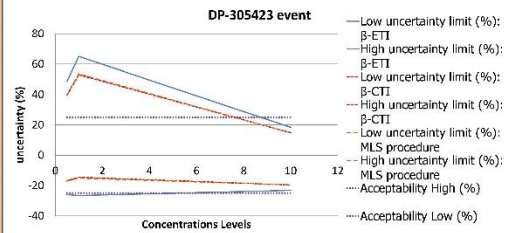


Fig. 1 Validation results of the real-time PCR method for the determination of GM content in GM DP-305423 event using three different accuracy profiles. (Acceptance limits $A = \pm 25\%$)

- 當以三種不同準確度曲線評估GM大豆DP-305423品項10%、1%、0.5%三個濃度於real-time PCR定量分析後的結果，並根據ENGL規定以此定量分析之真實性需為±25%的規範下設定可接受界限，結果顯示出在ENGL規定下評估此定量分析程序是否具長期有效性的表現時發現濃度越高其結果越穩定，低濃度的分析結果則顯示較不具有良好的長期有效性表現且其MU值也較高 (Fig. 1)。

D. 比較準確度曲線與IQC程序評估GM作物不同濃度之MU

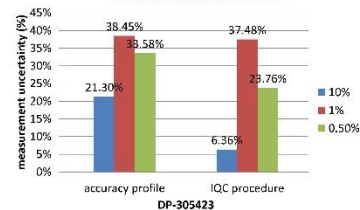


Fig. 2 Comparison of accuracy profile and IQC procedure in estimating MU of different concentration in GM soybean DP-305423 event.

- IQC程序是以top-down法為基礎，利用估算再現性，再加上偏差造成的變異來估計MU的一種評估方法。
- 本研究以準確度曲線和IQC程序評估GM大豆DP-305423品項在三種不同濃度下定量分析之結果時，發現皆顯示出一致的趨勢，且以準確度曲線評估之結果值皆大於IQC程序的結果值 (Fig. 2)。

E. 結論

- 目前GM作物定量分析中之MU的評估方法主要為bottom-up法，此法需要估計所有不確定度來源因子的變異，但並非所有因子的變異都容易被估計。
- top-down法則是利用評估分析方法的整體表現，並藉由計算再現性來評估MU，但不包含分析方法所造成的偏差不確定度。
- 本研究利用的IQC程序除了可以估計偏差的不確定度，使得top-down法能更加完備外，並可評估此定量分析方法是否符合目的適性 (fit for purpose)。
- 另外，本研究亦透過準確度曲線作為評估方法，研究結果發現β-ETI、β-CTI及MLS三種評估方法中所估計的最大MU值可達到以bottom-up法估計之效果，並可節省bottom-up法需估計所有不確定度來源的時間及降低人為的誤差，同時亦可評估定量分析方法是否符合目的適性。



利用零膨脹負二項模式模擬基因改造大油菜對小油菜之花粉飄散距離 — 以台中霧峰地區為例

楊芷翎 郭寶鐸
國立中興大學 農藝學系



研究目的

基改 (genetically modified, GM) 作物容易藉由花粉及種子傳播或是種子留存土壤的方式使得轉殖基因非預期出現 (adventitious presence) 在栽培 non-GM 作物的田區, 使其產生種子混雜或是產生帶有轉殖基因的子代, 造成基因流布 (gene flow), 對此經常會以在田區間設置隔離距離的方式因應 GM 作物與 non-GM 作物的共存問題, 並根據各個國家對於 GM 產品的接受程度設定不同的 GM 成分標示規範

本研究於台中霧峰地區進行大油菜對小油菜花粉飄散試驗, 蒐集得六期作的試驗資料並以零膨脹負二項模式 (zero-inflated negative binomial model, ZINB) 進行配適, 配適完成後分別討論在臺灣法規規定 3% 以及歐盟所訂定的 0.9% 標示門檻值推薦設置的隔離距離。

材料方法

(一) 試驗設計

本研究探討台中霧峰地區油菜花粉飄散情況, 田間試驗執行期間自 2013 年至 2015 年, 試驗地點於行政院農業委員會農業試驗所 (TARI) 以及國立中興大學農業暨自然資源學院試驗場 (AES) 執行。以大油菜品種德隆 18 號 (Deza No.18) 及小油菜品種農興 80 天 (Nonxing 80 days) 分別模擬 GM 對非 GM 油菜之花粉飄散, 田區配置將花粉貢獻小區設置於中段, 兩個田區均分別與花粉接受親小區接壤, 花粉貢獻親小區與花粉接受親小區之間不設置任何隔離距離或緩衝區, 藉此模擬在最嚴重的狀況下的花粉飄散情況 (Fig. 1)。

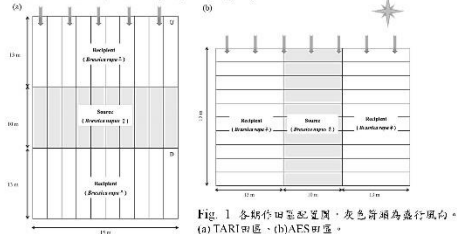


Fig. 1 各種作物田區配置圖, 灰色箭頭為盛行風向。
(a) TARI 田區, (b) AES 田區。

(二) 小油菜子代異交率調查

接受親小區中以距離花粉貢獻親小區 0.35 m 處作為採樣的起始點, 並以株距 0.7 m 為間隔設置採樣點, 單畦收穫 19 個不同距離的採樣點, 每一田區收穫兩側共 380 個採樣點 (Fig. 2)。自各採樣點中隨機選取 384 粒完整的種子播種於穴盤內, 每次播種三粒種子, 於幼苗二葉齡時調查子代之葉片外表型。具有異交情形的 F1 子代葉片可觀察到近似父本的外型 (Fig. 3); 非異交的子代葉片表現型如母本 (Fig. 4), 以觀察到的子代個體總數中具有異交情形的 F1 子代個體數比例做為樣本之異交率 (cross pollination rate, CP%)。

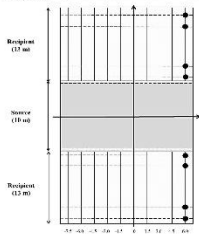


Fig. 2 採樣點位置標示圖。

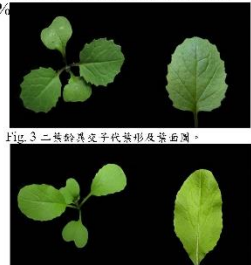


Fig. 3 二葉齡非異交子代葉形及葉面圖。
Fig. 4 二葉齡非異交子代葉形及葉面圖。

(三) ZINB 模式

ZINB 模式分為兩部分: (a) 目標事件的發生次數是否為零值的伯努利 (Bernoulli) 分布以及 (b) 描述目標事件的發生次數的負二項分布 (式(1))。

$$P(Y_i = y_i) = \begin{cases} \pi_i + (1 - \pi_i) \cdot g(y_i) & y_i = 0 \\ (1 - \pi_i) \cdot g(y_i) & y_i > 0 \end{cases} \quad (1)$$

$$g(y_i) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_i + 1)} \left(\frac{\alpha \mu_i}{1 + \alpha \mu_i} \right)^{\alpha^{-1}} \pi_i = \frac{\exp(Q \cdot \gamma(x, y))}{1 + \exp(Q \cdot \gamma(x, y))} \quad \mu_i = \exp(Q \cdot \gamma(x, y)) \quad (2)$$

於 ZINB 模式中代入 Nathan *et al.* (2012) 所整理的飄散核心函數 (表一) (dispersal kernel function, $\gamma(x, y)$) (式(2))。

引用文獻

Nathan, R., F. Klein, J.J. Robledo-Arnuncio, and F. Revilla. 2012. Dispersal kernels: review, p. 186–210. In Clobert, J., Baguette, M., Benton, T.G., Bullock, J.M., Dispersal Ecology and Evolution, Oxford University Press.

Table 1 Nathan *et al.* (2012) 所整理的飄散核心函數 $\gamma(x, y)$ 。

Kernel name	$\gamma(x, y)$ expression	Parameter values	Kernel name	$\gamma(x, y)$ expression	Parameter values
Gaussian	$\frac{1}{\sigma^2} \exp(-\frac{r^2}{\sigma^2})$	$a > 0$	Power law	$\frac{(b-2)(b-1)}{\pi^2 a^2} (1 + \frac{r^2}{a^2})^{-b}$	$a > 0; b > 2$
Neg. Exponential	$\frac{1}{2\sigma a} \exp(-\frac{r}{a})$	$a > 0$	Lognormal	$\frac{1}{(2\pi)^{1/2} b r^2} \exp(-\frac{(\log(r/a))^2}{2b^2})$	$a > 0; b > 0$
Exponential power	$\frac{b}{2\pi a^2 \Gamma(\frac{b}{2})} \exp(-\frac{r^b}{a^b})$	$a > 0; b > 0$	Weibull	$\frac{b}{2\pi a^2} r^{b-2} \exp(-\frac{r^b}{a^b})$	$a > 0; b > 0$
Logistic	$\frac{b}{2\pi a^2 \Gamma(\frac{b}{2})} (1 + \frac{r^b}{a^b})^{-1}$	$a > 0; b > 2$	Gamma	$\frac{1}{2\pi a^2 \Gamma(b)} (\frac{r}{a})^{b-2} \exp(-\frac{r}{a})$	$a > 0; b > 0$
WALD	$\frac{b}{\sqrt{2\pi r^3} \Gamma(\frac{b}{2})} \exp(-\frac{b(r-a)^2}{2a^2 r^2})$	$a > 0; b > 0$	Log-sech	$\frac{1}{(\pi^2 b r^2)} \frac{1}{(\frac{r}{a})^{b+1} + (\frac{r}{a})^{1-b}}$	$a > 0; b > 0$
2Dt	$\frac{(b-1)}{2\pi a^2} (1 + \frac{r^b}{a^b})^{-1}$	$a > 0; b > 1$			

分析步驟



結果與討論

Table 2 以 ZINB 模式分別配適 TARI 試驗資料及 AES 試驗資料之均方根誤差 (RMSE), 校正平方和 (Adj. R²) 以及赤池信息量 (AIC)。

Model	Training dataset			Testing dataset		
	RMSE	Adj R square	AIC	RMSE	Adj R square	AIC
Gaussian	0.008067	0.24561	-10984.66450	0.009569	0.32734	-10595.35296
Neg. Exponential	0.007896	0.27719	-11033.42555	0.009260	0.37034	-10669.98249
Exponential power	0.007852	0.28453	-11044.05593	0.009329	0.36046	-10651.22939
Logistic	0.007856	0.28381	-11042.90523	0.009398	0.35092	-10634.35465
WALD	0.007861	0.28293	-11041.51673	0.009370	0.35471	-10641.03909
2Dt	0.007849	0.28308	-11044.93386	0.009326	0.35952	-10649.55634
Power law	0.007909	0.27419	-11027.70564	0.009250	0.37116	-10670.46829
Lognormal	0.007855	0.28412	-11043.39916	0.009361	0.35602	-10643.35542
Weibull	0.007857	0.28373	-11042.77826	0.009317	0.36209	-10654.14132
Gamma	0.007850	0.28486	-11044.58032	0.009339	0.35902	-10648.66872
Log-sech	0.007850	0.28485	-11044.56908	0.009347	0.35800	-10646.86076

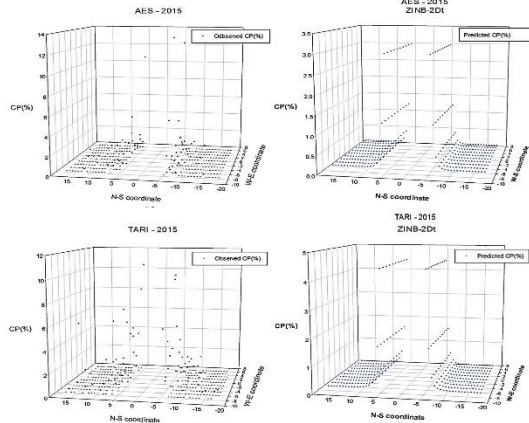


Fig. 5 2015 年 CP% 之 3D 散佈圖及以 ZINB-2Dt 模式配適之 CP% 之 3D 散佈圖。

Table 3 以 ZINB-2Dt 模式推算門檻值 0.9% 及 3% 之最佳隔離距離。

Model	Threshold of 0.9%		Threshold of 3%	
	Isolation distance (m)	Isolation distance (m)	Isolation distance (m)	Isolation distance (m)
ZINB-2Dt	1.78	0.78		

本研究在田間試驗中以大油菜及小油菜分別模擬 GM 對非 GM 油菜之花粉飄散, 並利用 ZINB 模式結合 $\gamma(x, y)$ 對資料進行配適, 根據各項模式評估準則以 ZINB-2Dt 模式表現最佳 (Table 2)。分別觀察 2015 年度 TARI 田區及 AES 田區模式預測結果與觀測值, 兩田區之異交率皆隨著距離增加而下降, 此外模式預測結果在距離貢獻親 1 m 內模式遞減速度快, 直至 3 m 後趨於平緩 (Fig. 5)。ZINB-2Dt 模式推薦的隔離距離在門檻值 0.9% 及 3% 下分別為 1.78 m 及 0.78 m (Table 3)。



不同區間估計方法在群組試驗中進行GM種子檢定的比較

蔡政煌 郭寶鈺
國立中興大學農藝學系

研究目的

種子生產過程須確認種子批 (seed lot) 中的基改種子比例是否能符合各國所規範之最低需求, 除了避免基改作物流入田間外, 也保障生產的種子品質。

本研究使用群組試驗 (group testing) 做為統計方法, 比較在相同的群組個數及群組內個別種子個數下比較Wald、Exact、貝氏等尾區間、最高後驗密度區間的覆蓋機率及區間期望寬度, 以利評估在不同情況下各種估計區間之表現。

材料與方法

資料來源

表 1. 以群組試驗檢定種子批中GM玉米比例之實例分析。

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Assumption						
Number of units by pools (n)	50	200	400	50	200	400
Limit of detection (q)	0.0005	0.0005	0.0005	0.0025	0.0025	0.0025
Concentration of AF per unit (d)	0.08	0.08	0.08	0.14	0.14	0.14
Minimum number of AF units required for detection (b)	1	2	3	1	4	8
Total number of tested pools (N)	100	100	100	100	100	100
Number of defective pools (y)	20	20	20	20	20	20
Estimates of the proportion of AF units in the population and their confidence intervals						
Maximum likelihood estimate (p)	0.0045	0.0041	0.0038	0.0045	0.012	0.0140
Two-sided 95% upper confidence limit	0.0069	0.0054	0.0047	0.0069	0.0136	0.0157
Two-sided 95% lower confidence limit	0.0027	0.0031	0.0031	0.0027	0.0096	0.0124
One-sided 95% upper confidence limit	0.0085	0.0052	0.0046	0.0065	0.0133	0.0154

(Montesinos-López *et al.*, 2010)

本研究之數據皆來自Montesinos-López *et al.* 於2010年所發表的文獻中, 針對GM玉米使用群組試驗來評估群組中個別種子為GM種子的機率 (表1)。群組內個別種子總數分別為50、200及400, 群組個數為100, 假設之陽性群組個數為20; 另一筆數據參考加拿大貿易出口的亞麻種子批之中, GM亞麻FP967在種子批中存在比例之文獻, 在研究中指出, FP967汙染亞麻種子批之比例約為0.8% (Vanella *et al.*, 2014), 並根據國際種子檢查規則中所規定, 亞麻於純淨度分析中最小之供試樣本重量為15克, 加拿大糧食委員會將供試樣本重量訂為60克, 相當於10000個亞麻種子(Canadian Grain Commission, 2018)。

群組試驗

群組試驗是將一定數量的樣本混合成數個群組, 接著再對群組進行試驗。若某一群組被檢驗出為陽性, 則該群組內至少含有一個以上的陽性樣本; 若某一群組被檢驗出為陰性, 則該群組內所有樣本皆為陰性。

一般而言, 若要使群組試驗達到節省試驗次數的效果, 單一樣本被檢驗出為陽性的機率需小於1/3 (Hwang, 1971), 而基改種子比例檢測的最低規範通常設定為10%以下, 符合使用群組試驗的條件。

貝氏統計方法

貝氏信賴區間係指使用貝氏事後分布建立的估計區間。貝氏統計學中將個別種子被檢出為基改種子之機率p視為一隨機變數, 假設p服從某一機率分布 $\pi(p)$, 並對概度函數 $f(y|p)$ 進行修正後計算出事後分布 $g(p|y)$ 。在顯著性水準為 α 下, 貝氏信賴區間使用事後分布之 $\alpha/2$ 百分位數與 $1-\alpha/2$ 百分位數做為 $100(1-\alpha)\%$ 貝氏信賴區間之上下限。

$$\int_{a/2}^{1-\alpha/2} g(p|y) dp \geq 1-\alpha$$

分析準則

1. 區間寬度

即一估計區間上限與下限相減所得之長度。相同信心水準 α 下, 若估計區間之寬度愈小, 涵蓋真實機率p之信心愈高。

2. 覆蓋機率

定義某一機率值 p_i 在給定的數值條件之下得出之區間 $CI(p_i)$, 則 $CI(p_i)$ 所能涵蓋 p_i 之機率稱為覆蓋機率。若某區間估計方法之覆蓋機率愈接近信心水準, 代表該估計區間表現愈佳。

參考文獻

- Hwang, F. K. 1971. A minimax procedure on group testing problems. *Tamkang J. Math.* 2: 39-44.
- Montesinos-López, O.A., A. Montesinos-López, J. Crossa, K. Eskridge and C.M. Hernández-Suárez. 2010. Sample size for detecting and estimating the proportion of transgenic plants with narrow confidence intervals. *Secd Sci. Res.* 20: 123-136.
- Vanella, R., A. Weston, P. Brodmann, and E. Kübler. 2014. Development of an event-specific assay for the qualitative and quantitative detection of the genetically modified flax CDC Triffid (FP967). *Food Control* 41: 128-133.
- Canadian Grain Commission (2018) Sampling and Testing Protocol for Canadian Flaxseed Exported to the European Union. Available at https://ec.europa.eu/food/sites/food/files/plant/docs/sc_modifigenet_20180319_sum_ann2.pdf (accessed March 2018)

結果與討論

表 2. 在信心水準為0.95下模擬區間估計方法之覆蓋機率與期望區間寬度。

n	y	Method	Coverage				
			Wald	Exact	Bayes equal-tailed	HFD	
p=0.05	1	Car. Pool (n=50) + 50	0.91276 ± 0.00887	0.99276 ± 0.00887	0.99276 ± 0.00887	0.99276 ± 0.00887	
		Exp. Wald	0.91817	0.99496	0.99379	0.99394	
	2	Car. Pool (n=200) + 200	0.91871 ± 0.00338	0.99371 ± 0.00338	0.99371 ± 0.00338	0.99371 ± 0.00338	
		Exp. Wald	0.91871	0.99396	0.99391	0.99412	
	p=0.08	10	Car. Pool (n=50) + 50	0.962135 ± 0.00556	0.99679 ± 0.00556	0.962135 ± 0.00556	0.962135 ± 0.00556
			Exp. Wald	0.96296	0.99741	0.96292	0.96314
20		Car. Pool (n=200) + 200	0.95204 ± 0.00352	0.99354 ± 0.00352	0.95204 ± 0.00352	0.95224 ± 0.00354	
		Exp. Wald	0.95291	0.99405	0.95285	0.95309	
p=0.15		10	Car. Pool (n=50) + 50	0.98128 ± 0.00309	0.99521 ± 0.00309	0.98128 ± 0.00309	0.98128 ± 0.00309
			Exp. Wald	0.98201	0.99552	0.98211	0.98229
	20	Car. Pool (n=200) + 200	0.98087 ± 0.00285	0.99625 ± 0.00285	0.98087 ± 0.00285	0.98087 ± 0.00285	
		Exp. Wald	0.98174	0.99675	0.98174	0.98194	
	p=0.20	10	Car. Pool (n=50) + 50	0.95865 ± 0.00318	0.99426 ± 0.00318	0.95861 ± 0.00308	0.95861 ± 0.00307
			Exp. Wald	0.95819	0.99501	0.95819	0.95841
20		Car. Pool (n=200) + 200	0.95725 ± 0.00279	0.99321 ± 0.00279	0.95726 ± 0.00268	0.95726 ± 0.00271	
		Exp. Wald	0.95725	0.99394	0.95725	0.95746	
p=0.30		10	Car. Pool (n=50) + 50	0.97948 ± 0.00278	0.99171 ± 0.00278	0.97946 ± 0.00269	0.97946 ± 0.00269
			Exp. Wald	0.97948	0.99239	0.97948	0.97969
	20	Car. Pool (n=200) + 200	0.98032 ± 0.00276	0.99059 ± 0.00272	0.98034 ± 0.00267	0.98034 ± 0.00270	
		Exp. Wald	0.98032	0.99102	0.98032	0.98053	
	p=0.40	10	Car. Pool (n=50) + 50	1.0	1.0	1.0	1.0
			Exp. Wald	1.0	1.0	1.0	1.0
20		Car. Pool (n=200) + 200	0.99998 ± 0.00001	0.99999 ± 0.00001	0.99999 ± 0.00001	0.99999 ± 0.00001	
		Exp. Wald	0.99999	0.99999	0.99999	0.99999	
p=0.50		10	Car. Pool (n=50) + 50	0.99998 ± 0.00001	0.99999 ± 0.00001	0.99999 ± 0.00001	0.99999 ± 0.00001
			Exp. Wald	0.99999	0.99999	0.99999	0.99999
	20	Car. Pool (n=200) + 200	0.99998 ± 0.00001	0.99999 ± 0.00001	0.99999 ± 0.00001	0.99999 ± 0.00001	
		Exp. Wald	0.99999	0.99999	0.99999	0.99999	

使用蒙特卡羅法模擬10000次檢定FP967存在於亞麻種子批中之比例, 並觀察四種區間估計方法在此檢定中所呈現出之覆蓋機率與期望區間寬度。在覆蓋機率的部份, 只有Exact區間的覆蓋機率在任何情況下皆維持在信心水準0.95之上, 符合其保守之特性, 其餘區間估計方法則在不同的群組個數之下呈現無規律性質之波動; 區間期望寬度的部份, 最高後驗密度區間在任何情況下皆具有最佳之區間期望寬度, 其餘區間估計方法則在群組個數的增加之下, 區間期望寬度的表現也隨之漸佳。

表 3. 考慮偵測極限下, 區間估計方法之區間期望寬度小於b下之機率大於水準 α 下所需的最小群組個數。

p	n	Exact		Wald	
		100	200	100	200
0.05	10	72	151	99	169
	20	57	89	49	59
	30	46	61	31	46
	40	37	43	23	37
	50	29	36	17	29
	60	23	31	13	23
0.08	10	81	152	101	170
	20	65	91	51	60
	30	52	64	38	52
	40	42	50	29	42
	50	35	43	23	35
	60	29	37	19	29
0.15	10	87	153	103	171
	20	70	92	53	61
	30	57	67	41	54
	40	47	52	33	47
	50	39	46	26	39
	60	32	40	21	32
0.20	10	91	154	104	172
	20	74	93	54	62
	30	61	69	43	56
	40	51	54	35	51
	50	43	48	28	43
	60	36	42	23	36
0.30	10	95	155	105	173
	20	78	94	55	63
	30	65	71	45	58
	40	55	57	37	55
	50	47	50	30	47
	60	40	44	24	40
0.40	10	98	156	106	174
	20	81	95	56	64
	30	68	73	47	60
	40	58	59	39	58
	50	50	52	32	50
	60	43	46	26	43

在偵測極限訂為0.0005、群組內個體數訂為50之下, 最高後驗密度區間之所需最小群組個數為四種區間估計方法之中最少, 表現也最佳; 在此實例分析之中, Exact區間與貝氏等尾區間之表現較為相等且基於最高後驗密度區間; 而Wald區間在考慮偵測極限之下表現為最差。

綜合表2及表3之結果, 在覆蓋機率表現之評估上, 除去具有保守特性之Exact區間, 其餘區間估計方法之覆蓋機率皆在信心水準附近呈現不具規則及趨勢之變動; 而區間寬度方面, 最高後驗密度區間在任何情況下皆具有最小之區間寬度表現, 因此本研究推薦在考慮偵測極限並使用群組試驗之下, 使用最高後驗密度區間將會得出最為精準之試驗結果。

The methods to prevent analysts unconsciously adjusting the result in the germination tests



Si-Ting Chen¹, Ciao-Yun Syu², Yu-Wen Kuo², Bo-Jein Kuo^{1*}

¹Department of Agronomy, National Chung-Hsing University (NCHU), Taichung, Taiwan

²Taiwan Seed Improvement and Propagation Station (TSIPS)

Introduction

Recently, many studies have discussed the under-dispersion in germination tests. Under-dispersion is that empirical variance between replicates is less than the binomial sampling variance. These studies concluded that analysts could unconsciously adjust the germination test results based on the thought that the seeds of the replicates were from the same seed lot. In order to prevent analysts unconsciously adjusting the germination test results, this study proposed two methods: the modified multiple-analyst tests (MMA tests) and multiple-investigation tests (MI tests).

Topic 1: MMA tests could effectively reduce the under-dispersion

1. Material and methods

- MMA tests were modified from the multiple-analyst tests^[1] for the small laboratories (Fig. 1).
- MMA data were generated by permuting the blind tests data.
- The permutation was repeated 10,000 times, so 10,000 MMA test results were obtained.
- There were 90 samples in a MMA test for rice (*Oryza sativa* L.) and pepper (*Capsicum annuum* L.), respectively.

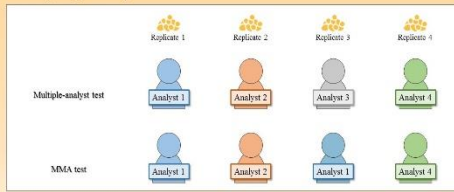


Fig. 1 The illustration of the multiple-analyst test and MMA test.

Topic 2: Under-dispersion was reduced with more investigation frequency

1. Material and methods

- In the MI test the investigation frequency was more than once.
- Three samples were randomly sampled from each seed lot and the investigation frequency of the samples was set to once, twice, and thrice, respectively (Fig. 2).
- The analyst was assigned randomly to each investigation and the results of replicates in all investigations were independent.

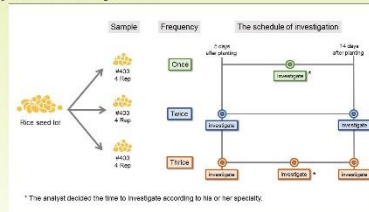


Fig. 2 The illustration of MI tests.

2. The evaluation of the under-dispersion

To evaluate the effect of MMA and MI tests on reducing the under-dispersion, F-values were calculated and the skewness of the distribution of F-value was used to evaluate the extent of under-dispersion. The F-value was the ratio of the empirical variance and the binomial sampling variance. The sk_{G1} , sk_{G2} , and sk_{G3} ^[2] were also calculated. When these coefficients closed to 1, it implied that the distribution of F-value was more right-skewed. It could be used to quantify the extent of under-dispersion.

3. Result

- For 5200 rice tests, the values of sk_{G1} , sk_{G2} , and sk_{G3} in MMA were all less than the values in single-analyst tests.
- For 9,917 pepper tests, the values were also less than the values in single-analyst tests.
- Fig. 3 showed that the effect of MMA on reducing the under-dispersion was more efficient for pepper than for rice.

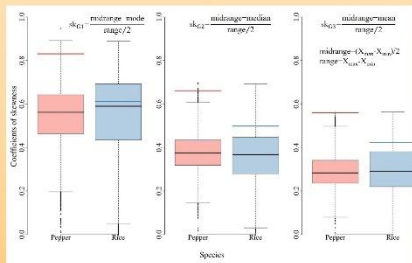


Fig. 3 The boxplots of sk_{G1} , sk_{G2} , and sk_{G3} for rice and pepper in MMA tests. The red lines denote the coefficients of skewness for rice in single-analyst tests. The blue lines denote the coefficients of skewness for pepper in single-analyst tests.

3. Result

- The sk_{G1} , sk_{G2} , and sk_{G3} were used to quantify the extent of under-dispersion due to the unconscious adjustment by analysts.
- The values of sk_{G1} , sk_{G2} , and sk_{G3} were decreased with more investigation frequency, meaning that fewer F-value value was less than 1 when the investigation frequency increased (Table 1).
- It indicated that the extent of the under-dispersion was reduced with more investigation frequency.

Table 1. The statistics of F-value for investigating once, twice, and thrice in MI tests.

Statistics	Frequency		
	Once	Twice	Thrice
Sample size	29	26	30
Standard deviation	0.8215	0.6430	0.8021
Midrange	1.8564	1.2162	1.4748
Mean	1.1965	0.9629	1.2268
Median	1.0298	0.8694	1.3292
Mode	0.8880	0.7060	1.0129
sk_{G1}	0.5475	0.4537	0.3223
sk_{G2}	0.4674	0.3084	0.1016
sk_{G3}	0.3731	0.2253	0.1730

Conclusion

The MMA tests reduced the extent of the under-dispersion due to the unconscious adjustment by analysts because of increasing the independence between the replicates. The MI tests also reduced the extent of under-dispersion by increasing the investigation frequency. Two methods can prevent analysts from adjusting the germination test results unconsciously and can effectively reduce the under-dispersion.

Reference

- Deplewski P. M. and M. Kruse. 2017. Online ring testing to elicit the cause of underdispersion in seed germination tests. *Crop Sci.* 57:3190-3202.
- Altınay G. 2016. A simple class of measures of skewness. Department of Econometrics, Bandirma Onyedi Eylul University, Turkey. 3pp.

研究室活動照片

