

林明傑、董子毅(2006)危險評估中 ROC 曲線在預測 2x2 表上與敏感度及特異度之關係 發表於第三屆測量、統計、方法學學術研討會 2006/09/23 於嘉義大學民雄校區 台北：台灣統計方法學學會

危險評估中 ROC 曲線在預測 2x2 表上與敏感度及特異度之關係

林明傑*

董子毅**

摘要

在預測犯罪時，經常會使用危險評估量表，而危險評估量表所使用的預測效度（又稱預測準確度）一般多是使用皮爾遜氏積差相關係數（Pearson's r ）與接收操作特徵曲線（Receiver Operative Characteristic Curve，簡稱 ROC）。然而，何謂 ROC？ROC 值所代表的意義為何？其在預測 2X2 表上與預測之敏感度與特異度有何關聯？犯罪防治人員應如何決定使用較好的量表？此均為本研究之目的。

從發展本土化的婚姻暴力危險評估量表及性侵害加害人靜態量表中發現 ROC 曲線值在預測 2X2 表(即是猜測會否再犯及實際是否再犯之表)上似乎是敏感度及特異度之平均值(即二者相加除以二)，經由本研究以公式及圖形之驗證，發現確如上述。並建議爾後所發展之危險評估量表均能清楚寫出再犯危險評估量表在預測 2X2 表上之 ROC 值及其敏感度與特異度，俾便量表之使用者及司法策者能對量表之準確度有進一步的正確了解。

*國立中正大學犯罪防治學系暨研究所專任助理教授

** 國立中正大學犯罪防治研究所碩士

本文之通訊作者為林明傑 電子郵件：crrmcl@ccu.edu.tw 通信地址：嘉義縣民雄鄉三興村大學路 168 號 國立中正大學犯罪防治學系暨研究所

壹、研究緣起

台灣的家庭暴力防治法於民國八十八年實施之後，各縣市家庭暴力防治中心即於民國九十年開始使用危險評估量表（Danger Assessment Scale, DA）（Campbell, 1995；林明傑、沈勝昂，2003）以評估及預測婚姻暴力加害人之暴力致命危險性。而自民國九十年起，國內亦開始使用美加之性罪犯再犯危險評估量表，如加拿大快速再犯危險評估表（Rapid Risk Assessment for Sex Offense Recidivism, RRASOR）（Hanson, 1997）、靜態因素九九評估表（Static-99）（Hanson & Thornton, 1999）、及明尼蘇達州之性罪犯評估表（Minnesota Sex Offender Screening Tool- Revised, MnSOST-R）（Epperson, Kaul, Huot, Hesselton, Alexander, & Goldman, 1997）（三量表均詳見林明傑，1999）進行對性罪犯的假釋評估及社區輔導與治療。

因此，可知專業人員經常會使用危險評估量表¹（risk assessment scale）以進行犯罪預測。而一般在評估危險評估量表之好壞的預測效度（又稱預測準確度）多是使用皮爾遜氏積差相關係數（Pearson's）與接收操作特徵曲線（Receiver Operative Characteristic Curve，簡稱 ROC）。然而，何謂 ROC？ROC 值所代表的意義為何？ROC 值是否與預測之敏感度與特異度有關聯？犯罪防治人員應如何決定使用較好的量表？上述均為犯罪學界較為陌生，且亦是本研究所欲探究之目的。

本文中所提及之名詞解釋如下：

1. 正陽性：原文為 **true positive**。指在犯罪預測中，猜測當事人會再犯，而其實際上也真的再犯。
2. 正陰性：原文為 **true negative**。指在犯罪預測中，猜測當事人不會再犯，而其實際上也不再犯。

¹ Bonta（1996）曾將危險評估區分為三代。第一代危險評估即以臨床人員依其臨床經驗預測未來犯行之可能性；而第二、三代危險評估則均是以精算方法進行危險評估，其均有研究基礎。第三代比第二代多納入理論與動態因素。詳見林明傑（2004）。

3.敏感度：原文為 *sensitivity*。指在犯罪預測中，表示在有再犯的群體中，猜中其會再犯的比率，此即為「正陽性率」，筆者認為亦可稱「正猜對率」使較能令人一目了然。

4.特異度：原文為 *specificity*。在犯罪預測中，表示在不再犯的群體中，猜測其不會再犯的比率，此即為「正陰性率」，筆者認為亦可稱「負猜對率」。

參、文獻分析

一、ROC 之介紹

ROC 曲線，原文名為 *Receiver Operative Characteristic Curve*，亦有名為 *Relative Operating Characteristic Curve*，可翻譯為「接收操作特徵曲線」或「相對操作特徵曲線」。

Hanson (1997) 曾指出可以用兩種方法來描述危險評估量表的準確度 (*predictive accuracy of the risk scale*)，即：(1) 相關係數 *r* (*correlation coefficient*)，(2) 在 ROC 曲線下之區域 (*area under the receiver operating characteristic curve*)。ROC 曲線最早是在無線電偵測技術及心理物理學中發展及運用，直到九〇年代中期才開始被用來預測暴力 (Quinsey, Harris, Rice, & Cormier, 1998)。ROC 是用來評估一個兩分變項中的分類表現，而在 ROC 曲線下之區域大小則可用來評估其預測準確度 (Mossman, 1994; Rice & Harris, 1995)。ROC 曲線上的每點將點繪出某測量法或某量表之每一次預測之正確率 (*hit rate*) 與錯誤率 (*false alarm rate*)，而該曲線下之分數則會分配在 .50 與 1.00 之間；其中，1.00 表示該預測係 100% 之正確，而 .50 則表示該預測不會比猜測來得準確，詳見圖 1²與圖 2¹。Hanson 與 Thornton (1999) 指出 ROC 比其他之預測方法 (如臨床人員之共識比率或相關係數) 好用，因為其並不受基線及選擇比率 (*base rates & selection ratios*) 之影響³。

註解 [U1]:

²本圖之對角線表 .50，即表示其係靠機會猜測 (*guess by chance*)。

³此基線係指再犯率基線或某型罪犯所占的比率基線；而選擇比率係指選擇某型罪犯之比率。

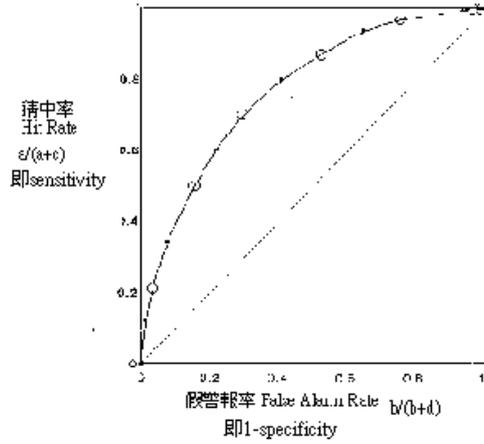


圖 1 ROC 曲線

- 註：1.Y 軸：為「猜中率」，即敏感度 (sensitivity)，亦即 $a/(a+c)$ 。X 軸：為「猜有實無 (假警報) 率」，即 1-特異度 (1-specificity)，亦即 $b/(b+d)$ ，圖中曲線之各點均顯示猜中率均大於假警報率。
- 2.計算 ROC 可在 SPSS10.0 中之統計圖 (graph) 下之工具列選單中尋得。
- 3.見 Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

		實際上是否再犯		
		確實再犯	確實不再犯	
預測 再 測 是 否	預 測 再 犯	Hits ; true positives 猜有實有 (正陽性) a	False alarm; false positives 猜有實無 (偽陽性) b (第二類型誤差)	a + b
	預 測 不 再 犯	Misses; false negatives 猜無實有 (偽陰性) c (第一類型誤差)	True negatives 猜無實無 (正陰性) d	c + d
		a + c	b + d	a + b + c + d

圖 2 預測實況分配圖

註：1.猜中率 (hit rate) = $a / (a + c)$ = 敏感度 (sensitivity)；假警報率 (false alarm rate) = $b / (b + d)$ = $1 - \text{特異度}$ (即 $1 - \text{specificity}$)。

2.陽性預測率 (positive predictive value, PPV) = $a / (a + b)$ ；陰性預測率 (negative predictive value, NPV) = $d / (c + d)$ 。

3 特異度高之量表，運用於盛行率低之族群時，大部分之陽性結果是假陽性。敏感度高之量表，運用於盛行率高之族群時，大部分之陰性結果是假陰性，即陽性與陰性結果之解釋，須考慮不同之場合與盛行率 (余光輝，2005；吳雪玉、廖華芳、姚開屏、李旺祚、王天苗、謝正宜，2005)。

二、選取切分點之模式與正陽性、正陰性之關係

從圖一的直向來看，實際只有兩群人，即為確實再犯者及確實不再犯者，若將每一群視為 100% 來看，只有兩種猜對率，即是正陽性之猜對率及正陰性之猜對率。因此，在評估任何危險評估量表是否有其預測的準確度，實需了解量表在敏感度與特異度的猜對率。在有些情形下，只須重視「正猜對率」(敏感度)要高，而有些情形則只須重視「負猜對率」(特異度)要高。但要如何考慮則可參考 Swets (1996) 之說法。

Swets (1996) 提出設定切分點 (或稱臨界值) 的三種取向：(1) 工業品質

管制取向 (industrial quality-control approach) 認為「猜有實無」之猜錯率是很嚴重且要避免的，其希望要將「偽陽性」(FP)降低至 5% 至 1%，因此其重點在需要將「正陰性」(TN)升高，而不考慮「正陽性」(TP)。(2) 工程失敗安全取向 (engineering fail-safe approach) 認定「猜無實有」之猜錯率是很嚴重且要避免的，且希望要將「正陽性」(TP)升高，而不考慮到「正陰性」(TN)。(3) 對等取向 (symmetrical approach) 乃認為兩種猜錯率均同樣嚴重，有必要均等考量 (包含均等考量正負猜對率之成本與效益) 並予以避免，因此此取向會將整體之猜對率 (包含正負猜對率) 拉到最高，而不會只及顧任何一方。

從 Swets (1996) 的說法中，可知無論選擇哪一種方法，都必須事先仔細評估評量者所重視的是什麼。筆者認為對於犯罪者之再犯或致命傷害預測，應以均衡考量加害者與受害者之人權為原則，給予平等的考慮，除非有很明顯之事實或研究顯示加害人或被害人會因為預測而造成負面傷害之成本效益明顯地偏向任何一方。因此，筆者建議犯罪預測之預測取向應以對等取向為原則來選擇量表之切分點，也就是平等地將正陽性與正陰性提到最高。

三、從發展本土量表中的發現

筆者從發展本土化之婚姻暴力危險評估量表 (Danger Assessment Scale, DA) (林明傑、張詩吟, 2004; 林明傑、史玉山、簡蕾如, 2003) 與台灣性罪犯靜態再犯危險評估量表 (林明傑、董子毅, 2005, 見表 1) 中，發現 ROC 在 2X2 表之表現時，其 ROC 值與敏感度及特異度之平均值 (即二者相加除以二) 相同。然而在檢索所有中外文獻之後，至今並未發現有討論 ROC 與敏感度及特異度之統計關係，因此筆者嘗試進一步進行統計公式之驗證。

肆、統計驗證

由於 ROC 曲線在 2X2 表時只會有一個轉折點，今假設 $Y = \text{sensitivity}$ (即垂直軸)， $X = \text{specificity}$ (而 $1-X$ 即是水平軸)，且 $0 < X, Y < 1$ ，則 y 即表示該點在「敏感度」(即垂直軸)之值，而 $1-x$ 即表示該點在「1-特異度」(即水平軸)之值(見圖 3)。該點之 ROC 曲線下的面積值即是「上三角形+左三角形+長方形」之總合，經由下式之演算可知在 2X2 之預測表上的 ROC 值就是敏感度與特異度(即正猜對率與負猜對率)之平均值。

統計之推算如下，並如圖 3。

假設：(1) $Y = \text{sensitivity}$ (即垂直軸)；(2) $1-X = 1-\text{specificity}$ (即水平軸)

長方形 = xy

上三角形 = $(1-x)y/2$

左三角形 = $(1-y)x/2$

$$\begin{aligned} \text{ROC} &= \text{長方形} + \text{上三角形} + \text{左三角形} \\ &= xy + (1-x)y/2 + (1-y)x/2 \\ &= xy + (y-xy+x-xy)/2 \\ &= (x+y)/2 \end{aligned}$$

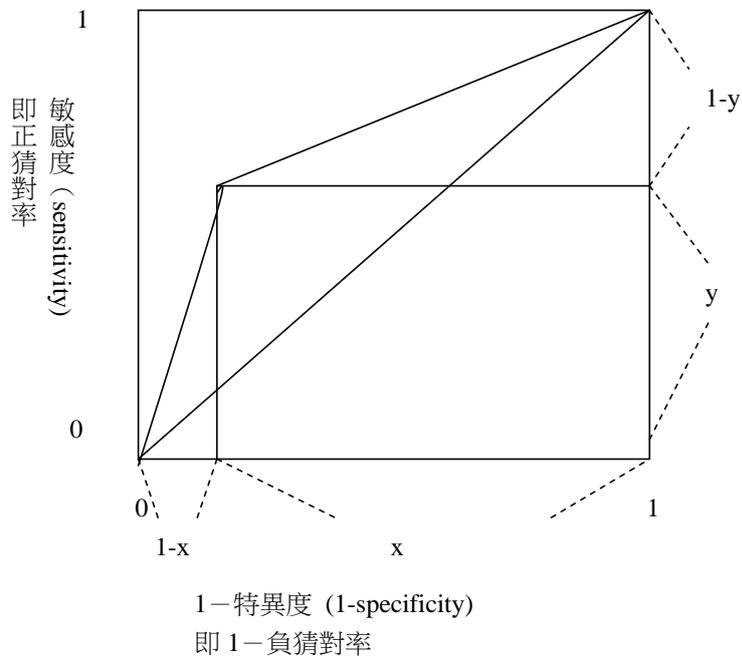


圖 3 ROC 曲線在 2X2 預測表上正負猜對率之關係

伍、結論

根據上述之公式驗證，可知在 2X2 之預測表上的 ROC 值就是敏感度與特異度（即正猜對率與負猜對率）之平均值。然而，也必須了解的是，如果所提供的 ROC 不是在 2X2 表上，而是以原始分數預測再犯危險，或以高中低之量表分數預測再犯危險，則其 ROC 預測準度多數上應會比只區分是否會再犯之危險分級後預測準確度之情形高；而後者若能選擇好的切分點，則可將 ROC 值之損耗到最小。

筆者建議日後之危險評估量表均能附上其預測在 2X2 表之表現，亦即能清楚寫出量表在 2X2 表之 ROC 值及正負猜對率，俾便使用量表之臨床評估者及刑事司法決策者能對該量表對再犯預測之準確度多少及正負預測之猜對率與猜錯

率均有所了解。

參考書目

中文資料

余光輝

- 2005 《實證醫學講義》2005/8/30 下載于
<http://www.cgmh.org.tw/intr/intr2/ebmlink/doc>

吳雪玉、廖華芳、姚開屏、李旺祚、王天苗、謝正宜

- 2005 〈"嬰幼兒綜合發展測驗"動作分測驗與"皮巴迪動作發展量表第二版"的診斷準確度〉。《台灣醫學》9(3): 312-322。

林明傑、董子毅

- 2005 〈台灣性罪犯靜態再犯危險評估量表(TSOSRAS)之建立及其外在效度之研究〉。《亞洲家庭暴力與性侵害期刊》1(1): 49-110。

林明傑、張詩吟

- 2004 〈婚姻暴力加害人再犯危險評估量表建立研究〉。《內政部委託研究計劃》。台北：內政部。

林明傑

- 2004 〈危險評估方法學〉。收錄於林明傑、沈勝昂編《法律犯罪心理學》373-387。台北：雙葉書廊。

林明傑 沈勝昂

- 2003 〈我國婚姻暴力加害人之危險評估：DA 量表在我國適用之研究〉。《犯罪學期刊》6(2): 177-216。

林明傑、史玉山、簡蕾如

- 2003 〈婚姻暴力加害人再犯危險與致命危險評估量表之研究〉。《內政部委託研究計劃》。台北：內政部。

英文資料

Campbell, J. C.

- 1995 Prediction of homicide of and by battered women. In J. C. Campbell (Ed.), *Assessing dangerousness: Violence by sexual offenders, batterers, and child abusers* (pp96-113). Thousand Oak, CA: Sage.

Epperson, D. L., Kaul, J. D., Huot, S. J., Hesselton, D., Alexander, W., & Goldman, R.

- 1997 Minnesota Sex Offender Screening Tool - Revised [Online] available
http://www.psychology.iastate.edu/faculty/epperson/mnsost_download.htm

Hanson, K. & Thornton, D.

- 1999 *Static 99: Improving actuarial risk assessments for sex offenders*. Ottawa, Canada: Department of the Solicitor General Canada. [Online] available
http://www.sgc.gc.ca/publications/corrections/199902_e.pdf

Hanson, K.

- 1997 The Development of a Brief Actuarial Risk Scale for Sexual Offense

Recidivism. Ottawa, Canada: Department of the Solicitor General Canada.

Online available at

http://www.sgc.gc.ca/publications/corrections/199704_e.pdf

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A.

1998 *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Swets, J. A.

1996 *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Mahwah, NJ: Lawrence Erlbaum Association.

The Relationship between ROC and Sensitivity and Specificity in 2X2 Table among Risk Assessment

Min-chieh Jay Lin⁺
Tzu-yi Dung

ABSTRACT

The risk assessment scale is usually used to predict the recidivism. It was always used the Pearson's r and Receiver Operative Characteristic Curve (ROC) to assess the predictive accuracy for the risk assessment scale. But what is the ROC, and what it means? How does the ROC relate to sensitivity and specificity? All the above were the topic of this study.

During the development of risk assessment scale in Taiwan, it was always found that the ROC seemingly is the mean of sensitivity and specificity, while using 2 by 2 table of "guess yes or not" and "reoffense or not". In this study, it was verified the previous notion through algebra. The authors suggested that it has to mark the values of ROC, sensitivity, and specificity under the 2 by 2 predictive table in the future risk assessment scale to make forensic practitioners and criminal justice practitioners have a better understanding on the hit rates and false rates of the risk assessment scale.

KEYWORDS: risk assessment, ROC, sensitivity, specificity

⁺ Min-chieh Jay Lin serves as an assistant professor in Department of Criminology, National Chungcheng University, Taiwan. Tzu-yi Dong earned masters degree in the same program as Lin. All correspondence should go to Min-chieh Jay Lin, Department of Criminology, National Chungcheng University, 168 University Road, Min-shiung, Chia-yi, Taiwan. Email address: crmmcl@ccu.edu.tw