防空警報響不響:Hypothesis Testing

last modified August 15, 2008

雷達系統是用來偵測空中的飛行物體。作戰的時候, 一旦發現敵機來襲, 必須隨即啓動警報系統, 一來備戰, 二來告知人員疏散。雷達系統不是人的眼睛, 如何判斷是否有飛行物體侵入? 何況人眼所及的距離有限, 爲有充分的時間因應, 還是要仰賴雷達系統更長距離的偵測能力。二十世紀初統計學家建立的假設檢定的機制, 成爲後來蓬勃發展的國防、太空及無線通訊工程的開端。

1 假設與決定

先來簡單描述雷達系統的運作; 雷達會發射電磁波到空中, 當電磁波在空中碰到阻礙 (譬如飛行物體), 將如光線碰到牆壁般的折射回來, 一部份的折射電波會回到雷達, 由雷達的接收器接收進來。接收進來的電波訊號經適當的取樣 (Sampling) 後成爲數據, 進行資料的分析與判斷。由於接收器是不停歇的接收訊號並轉換爲數據, 對於訊號的可能的來源先要有概念, 才能決定恰當的數學模式, 雷達接收的訊號大概來自:

1. 折射自飛行物體的電波。
2. 折射自空中阻體的微弱電波, 空中阻體如: 雲層、雨、水蒸氣、離子、風...。
3. 雷達系統本身的電子設備的干擾。
4. 環境的電磁波 (廣播、電視...)。

除第一項外, 其餘都可視爲互相獨立的干擾項, 其強度比折射自飛行物的電磁波相對小很多, 一般以隨機的雜訊 (Random noise) 來對待。統計學家的工作便從處理這些

接收轉換而來的數據開始, 希望透過這些數據來判斷訊號是否來自上述的第一項來源。統計學家做了兩個假設:

> H_0: no target present
> H_a: target present

其中 H_0 稱爲「原始假設 (Null hypothesis),」H_a 爲「對立假設 (Alternative hypothesis)。」[1] 以數學模式來表達, 可以寫成

$$H_0 : \mathrm{y} = \epsilon$$
$$H_a : \mathrm{y} = 3 + \epsilon$$

其中 y 代表接收訊號的變數, 在此並假設雷達電磁波的強度爲 3(單位), 變數 ϵ 則是所有雜訊的總和。根據中央極限定理, 變數 ϵ 將服從常態分配,[2]爲方便起見, 在此假設爲標準常態, 即 $\epsilon \sim N(0,1)$。ϵ 的常態分配假設非常重要, 這使得接收的訊號 y 雖具不確定性, 但卻服從某種已知的規律, 這樣便將問題帶進統計學家熟悉的領域。

> 問題:
> 根據一個接收訊號的樣本 y, 判斷兩個假設 H_0,H_a 之中, 何者爲眞。

這樣的文字敍述在統計學家的腦袋裡會變成談論機率的語言:

> 抉擇:
> 選擇 H_0, 如果 $P(H_0|y) > P(H_a|y)$, 否則, 選擇 H_a 。

上述文中的條件機率又稱爲「後驗機率 (a posteriori pobabilities)。」接下來的問題是, 這些後驗機率該如何計算呢? 當你讀到這裡時, 最好停下來想一想, 別急著往下看, 否則容易被作者帶壞了, 減弱了自己的思考能力。

> 想一想:
> 如果 $y = 4$, 比較可能是來自哪一個假設 (H_0 或 H_a) 的結果? 如果 $y = 2$ 呢? $y = -1$ 又如何?

[1]我們可以將原始假設與對立假設的內容互換嗎? 答案是不行。這事關重大, 說來話長, 讀者先接受這樣的安排, 以後再慢慢去找尋原因。

[2]簡單的說, 許多互相獨立的變數和將接近常態分佈。

不妨試著將問題倒過來想, 在 H_0 的假設成立下, 發生 $y = 4$ 的機率有多大? 如果是在 H_a 的假設下呢? 換句話說, 我們希望能計算並比較下列兩個機率:[3]

$$P(y|H_0), P(y|H_a)$$

這兩個機率也有個響叮噹的名號, 稱爲概似函數 (Likelihood function), 爲方便起見, 改寫爲 $p_0(y), p_a(y)$ 。根據 H_0, H_a的假設, 變數 y 服從常態分配, 其機率密度函數分別爲 $p_0(y), p_a(y)$, 寫成

$$\begin{aligned} p_0(y) &= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \\ p_a(y) &= \frac{1}{\sqrt{2\pi}} e^{-(y-3)^2/2} \end{aligned} \tag{1}$$

圖1展示這兩個概似函數的分佈情況。其中的三條虛線標示 $y = 4, 2, -1$ 等三個位置, 想試著回答前面的問題。從圖中很清楚的得出三個結果

$$\begin{gathered} \text{當 } y = 4 \text{ , } p_a(y) > p_0(y) \Longrightarrow \text{ 選擇 } H_a \\ \text{當 } y = 2 \text{ , } p_a(y) > p_0(y) \Longrightarrow \text{ 選擇 } H_a \\ \text{當 } y = -1 \text{ , } p_a(y) < p_0(y) \Longrightarrow \text{ 選擇 } H_0 \end{gathered}$$

由於常態假設的關係, 進一步分析 $p_a(y), p_0(y)$ 的比例關係, 得出

$$\lambda(y) = \frac{p_a(y)}{p_0(y)} = e^{y-1.5}$$

其中 $\lambda(y)$ 稱爲概似比 (Likelihood ratio)。[4] 於是我們的抉擇變爲, 當 $\lambda(y) \geq 1$ (或 $y \geq 1.5$) 時, 選擇 H_a , 否則, 選擇 H_0 。上述概似比可以進一步簡化爲

抉擇:

如果 $y \geq 1.5$, 選擇 H_a , 否則, 選擇 H_0 。

[3]請細心揣摩這種機率語言: 從文字的敍述(需求) 到條件機率的表達。

[4]條件機率的比例是個老把戲, 很多地方都派得上用場, 也常能得到很漂亮的結果, 希望初學者能好好品味賞玩, 至少多看幾眼。

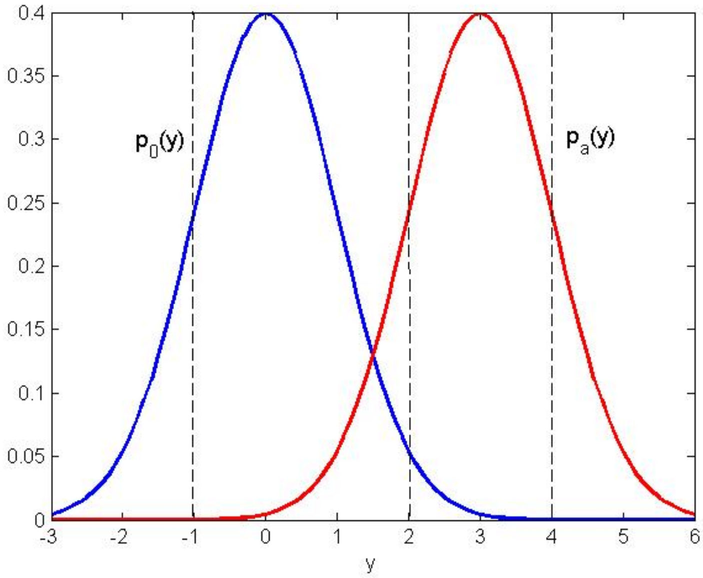

圖 1: 訊號 y 的機率密度函數。

這個結果直接從樣本值的大小下判斷, 此時藉以判斷的主角 y 稱為檢定統計量 (Test statistic), 這個檢定方法稱為概似比檢定 (Likelihood ratio test)。
推演至此, 看似順暢, 其實別忘了剛開始的用意是比較後驗機率, 而非概似函數。重述原來的想法:

> 抉擇:
> 如果 $P(H_a|y) \geq P(H_0|y)$, 選擇 H_a , 否則, 選擇 H_0 。

因為後驗機率通常比較不清楚, 之前才試著從概似函數推演回來, 這其中漏掉一個重要的東西。我們借用貝氏定理來還原事情的原貌, 列出後驗機率與概似函數的關係:

$$\begin{aligned} P(H_0|y) &= \frac{P(y|H_0)P(H_0)}{P(y)} = \frac{p_0(y)P(H_0)}{P(y)} \\ P(H_a|y) &= \frac{P(y|H_a)P(H_a)}{P(y)} = \frac{p_a(y)P(H_a)}{P(y)} \end{aligned} \tag{2}$$

後驗機率的比為

$$\frac{P(H_a|y)}{P(H_0|y)} = \frac{p_a(y)(1-P(H_0))}{p_0(y)P(H_0)} \geq 1 \tag{3}$$

其中 $P(H_a)$ 以 $1-P(H_0)$ 代入。上式以概似比表示寫爲

$$\lambda(y) = \frac{p_a(y)}{p_0(y)} \geq \frac{P(H_0)}{1-P(H_0)} \tag{4}$$

當考慮概似函數爲常態的假設後, 概似比檢定寫成 (作業 1)

> 抉擇:
>
> 如果 $y \geq 1.5 + \ln \frac{P(H_0)}{1-P(H_0)}$, 選擇 H_a , 否則, 選擇 H_0 。

這裡納入了當初遺忘的 $P(H_0)$, 也就是原始假設發生的機率。這個機率是多少呢? 能計算嗎? 先不管這些, 這個機率也有個名號, 稱爲「先驗機率 (a priori probability)。」通常先驗機率可以從已知的資料去估計, 或是假設毫無概念的均等分配。在這裡如果假設原始假設與對立假設發生的機會均等, 即 $P(H_0) = 0.5$, 則

$$\ln \frac{P(H_0)}{1-P(H_0)} = 0$$

概似比檢定又回到最初未考慮 $P(H_0)$ 的情況。

2 False Alarm 與 Probability of Detection

上一節介紹的概似比檢定法則只考慮到樣本發生的機率, 但機率與事實間的落差, 導致純以機率作爲判斷的依據可能會犯兩種錯誤。第一, 在原始假設 H_0 爲眞的情況下, 選擇了 H_a(或說拒絕了 H_0) , 稱爲型一誤差 (Type I error)。以雷達系統爲例, 雷達警報響了,[5]但事實卻是沒有敵機來襲 (H_0 爲眞), 一般稱爲 False Alarm。第二, 在對立假設 H_a 爲眞的情況下, 選擇了 H_0 (或說沒有拒絕 H_0), 一般稱爲型二誤差 (Type II error)。以雷達系統爲例, 警報器沒有響, 敵機的炸彈已經落下來了。圖2上的兩塊面積分別代表發生型一誤差與型二誤差的機率。

假設檢定的專有名詞「型一誤差」常被稱爲 False Alarm, 其源由便是從雷達系統而來。[6]至於型二誤差的機率, 一般較少直接引用, 反而重視它的另一面, 即在對立假設

[5]根據樣本 y , H_a 發生的機率較高, 於是判斷敵機來襲, 拉起警報。

[6]這裡可以稍解釋爲何原始假設不能隨意給定, 才能維持名符其實的 False Alarm。

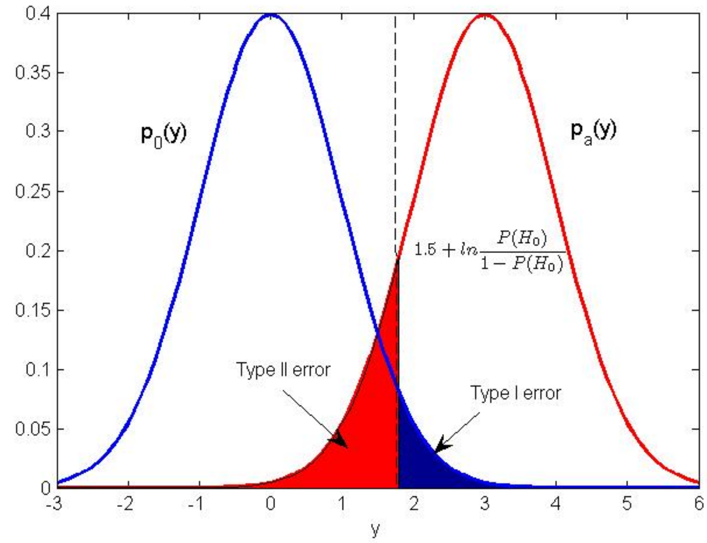

圖 2: 型一與型二誤差機率。

H_a 爲眞的情況下, 正確選擇了 H_a(也就是拒絕 H_0), 在雷達系統裡稱爲偵測機率 Probability of detection, 在一般統計應用上則稱爲檢定力 (Power of the test)。這麼說來, 我們是否能做些什麼以降低 False Alarm 的機率 (型一誤差機率), 另一方面提升偵測機率 (檢定力), 達成雙贏的局面? 圖2否定了這個可能性, 當虛線往右移動時, 型一誤差變小, 但同時增大了型二誤差 (即降低了檢定力), 反之增強了檢定力, 型一誤差機率也跟著變大。

高偵測機率在雷達系統裡非常重要的, 我們不希望看到當敵機來襲時, 人們還在逛大街。倒是 False Alarm 的錯誤在雷達系統裡相對輕微, 因爲錯誤的警報造成的傷害較小, 頂多引起民怨與不信任感罷了, 因爲當警報響起時, 不爽的人還是會跑。

3 作業

1. 從式 (4) 推展出概似比檢定的準則爲

$$y \geq 1.5 + \ln \frac{P(H_0)}{1 - P(H_0)}$$

2. 假設 $P(H_0) = 0.5$, 計算圖2的型一誤差與型二誤差機率。

3. 從第一節的結果得知, 當概似比檢定設定爲 $\lambda(y) \geq 1$ 時, 其型一誤差與檢定力就已經固定下來, 沒有第二節分析的提高檢定力的空間。但如果將概似比檢定設定爲 $\lambda(y) \geq \eta$, $\eta > 0$, 在固定型一誤差機率爲某個特定值 α 的情況下, 概似比檢定可以得到比任何檢定都大的檢定力, 稱爲 Neyman-Pearson 定理。假設 $P(H_0) = 0.5, \alpha = 0.1$, 試計算對應的 η 及檢定力。

4. 第一節的分析只針對一個樣本 y, 現擴大爲 n 個樣本, $y_1, y_2, \cdots, y_n$, 請改寫概似比檢定並計算型一誤差機率與檢定力。

參考文獻

[1] Anthony D. Whalen, "Detection of Signal in Noise,"Academic Press。