

第五章 假設檢定的關鍵概念

Jih-Chang Yu

2023年8月31日

動機範例1

全國教育統計中心報告2007年四年級閱讀分數。

全國平均值：220.99，標準差：35.73。

數據來自全國教育進步評估。

假設：您的學區在閱讀教學方面表現出色。

樣本數據與初步推論

隨機選取了您學區的50名四年級學生。

進行了相同的考試。

樣本平均值：230.2。

似乎支持您的假設。

批評者的擔憂：樣本的運氣可能影響結果。

消除樣本變異性

資源有限，只能使用樣本數據。

問題：測試所有四年級學生是否會產生類似於全國平均值的均值？

解決樣本變異性作為解釋的問題。

統計推論的組成部分

統計推論包括兩個部分：

關於參數值的陳述。

陳述可靠性的衡量，通常以概率形式表示。

傳統推論目標：

進行假設檢定：

假設具體的參數值或關係。

基於樣本統計量做出決策。

可靠性：錯誤決策的概率。

估計參數：

使用樣本統計量，通常是區間。

可靠性：區間的置信度。

動機範例2

公司在8盎司罐中包裝鹽焗花生。

目標：控制機器中罐子的花生量。

功效定義為平均每罐8盎司。

避免過度或不足填充。

樣本與假設檢定

使用16個罐子監控控制。

在不同時間間隔隨機抽取樣品罐。

測量每個樣品罐的重量。

使用這16個罐子的平均重量進行假設檢定。

假設：機器正常運作。

昂貴的調整

如果機器不正常運作：

需要進行昂貴的調整。

零假設的定義

定義：零假設是關於一個或多個參數值的陳述。

該假設代表現狀，通常除非樣本結果強烈暗示其為假，否則不會被拒絕。

零假設 (H_0)：無效效果、無差異或無關係。

備擇假設的定義

定義：備擇假設是一個與零假設相矛盾的陳述。

當零假設被拒絕時，該假設被接受。

備擇假設通常稱為研究假設，因為它通常意味著要執行某些行動、花費一些資金或推翻某些既定理論。

備擇假設 (H_1 或 H_a)：預期效果、差異或關係。

拒絕區域的定義

定義：拒絕區域是由臨界值確定的一個值域。

它代表在假設檢定中會導致拒絕零假設的樣本結果集合。

落在拒絕區域內的樣本結果提供了反對零假設的強有力證據。

假設檢定中的拒絕區域

考慮包裝鹽焗花生的例子：

零假設 (H_0) 假設機器正常運作，即均值為8盎司。

備擇假設 (H_1) 提示機器未正常運作，導致偏離8盎司。

為了決定是否拒絕 H_0 ，我們設置了一個臨界區域。

在假設檢定中，我們設置一個臨界值，用 C 表示。臨界值 C 定義了拒絕區域的邊界。

如果檢驗統計量 T 大於或小於 C ，我們拒絕零假設 (H_0)。

I型和II型錯誤

I型錯誤：當實際為真時錯誤地拒絕 H_0 。

這類似於“假陽性”情況。

就像指控一個人做了他沒做的事。

I型錯誤的概率用 α 表示。

II型錯誤：當實際為假時未能拒絕 H_0 。

這是一種“假陰性”情況。

就像錯過了眼前的東西。

II型錯誤的概率用 β 表示。

顯著性水平

顯著性水平 (α) 設定了拒絕零假設的門檻。

常見值：0.05, 0.01。

例如：設置 $\alpha = 0.05$ ，這意味著我們願意接受5%的概率來犯 I型錯誤（當 H_0 為真時拒絕它）。