

統計推論與結語

主要資料來源：《統計與真理》，C. R. Rao 著(美)，科學出版社，2004。

所謂的統計推論數據分析，就是基於一指定的隨機概率模型來估計未知參數，進行相應的假設檢驗，預測未來的觀測值，以及做出決策……等的統計方法。所選擇的模型並不要求能解釋全部觀測所得的數據，僅需對指定的問題提供有效的回答即可。為了了解給定數據的性質，要進行更廣泛的數據分析，以便發現所擁有的數據能回答哪些問題，從而提出新問題和計劃進一步的調查研究。我們也應該探索在同一組數據下，利用不同模型出現的不同結果，來回答不同問題的可能性。

統計處理必須對所收集的數據作一定的檢驗，如果數據具有某些特殊的特徵時，必須在處理過程中考慮，而且在統計處理中還要不斷監視整個過程，以決定是否需要對原定的處理做出修改。

統計分析的目的則是「從觀測得到的數據中，提取所有的訊息。」記錄的數據有時存在缺陷，例如記錄誤差和異常值，有時甚至可能是偽造的，在分析前應詳細考察交叉驗證，以便發現可能存在的缺陷並了解數據的特徵。下一步則是提出合理的隨機概率模型，進行數據推論分析，包括未知參數的估計、假設檢定、對未來觀測值的預報及做出決策。數據分析也會對提出新問題，和計劃進一步的調查研究提供訊息。

統計學家和實驗科學家需要合作研究。一個統計學家可以幫助科學家設計有效的實驗，在科學家提出的問題上獲得最多訊息，從而使科學家能檢測自己提出的假說，並在數據產生矛盾跡象時進行修改。實驗設計之父費雪就指出：

實驗結束後，向一個統計學家諮詢的常常是，要他提出一個後續的檢驗，他或許能指出實驗失敗的原因。