

單元 30: 一些一般性註解 (課本 §4.8)

一. 密度函數 (density function) 為自然界中真實資料母體的理論模型 (theoretical model).

二. 兩個問題:

(1) 如何選擇模型?

(2) 選錯模型時, 影響如何?

(2) 的回答:

(i) 機率模型的目的是在提供一種機制, 由此可根據樣本中的資訊對母體做推論.

(ii) 觀察到樣本的機率 (或某種比率) 對母體的推論是有幫助的.

(iii) 對母體頻率分布而言，不好的配適度的密度函數可能（但不必然）會產生不正確的機率敘述，而導致有關母體的錯誤推論。

(iv) 好的模型是指能產生好的母體推論的模型。

(1) 的回答：

(i) 選擇好的模型有時是一種理論上的考量，如，根據事件在時間上的隨機行為，卜松隨機變數會是適宜的，因此，接連二事件的間隔時間就會是一種指數分布；同理，若 a, b 均為整數且 $a < b$ 則第 a 次事件與第 b 次事件的間隔時間就會是 $\alpha = b - a$ 的 gamma 分布。另外，中央極限定理也會列出一些資料分布是符合常態分布近似的條件。

(ii) 第二種選擇模型的方法是建構由母體所得資料的頻率直方圖，並選擇在視覺上最類似頻率曲線的密度函數。

(iii) 並不是所有模型的選擇都是完全的主觀：有可用的統計過程（程序）來檢定“一母體頻率分布是某一特

殊類型”的假設，也可計算出數種分布的配適度量度並選取最好的。許多常用的統計推論方法的探討（研究）也已發展出，可用來決定由於不正確母體模型而引發的錯誤推論的程度。另人欣慰的是，許多推論的統計方法對於“基本母體頻率分布的型式”的假設是不敏感的 (insensitive)。

(iv) 均勻，常態，gamma, beta 以及 Weibull 分布提供多種配適於許多母體頻率分布的密度函數。