

Graduation

Assignment 3, Due Nov. 29/2024

Fall 2024

1. 本題目的在於測試如何選擇核修勻及 Spline 的參數。請依據 2023 年台北市/高雄市人口資料，包括死亡觀察值及簡易生命表：
 - (a) 測試核修勻的不同環寬 (Bandwidth) 及核函數 (Kernel function) 選擇，先說明參數選取標準，再選取可以接受的參數值，並根據結果建議如何選擇核修勻的參數。
 - (b) 測試 Spline 修勻不同參數 (例如：節點數及其年齡) 的修勻效果，先說明參數選取標準，並根據修勻結果建議可以接受的參數值。
 - (c) 測試可加性模型 (Generalized Additive Model) 可否用於死亡率修勻，請以本題的 2023 年台北市/高雄市死亡率資料為依據。

2. 假設資料來自於

$$y_i = \sin(2\pi x_i) + \varepsilon_i, \varepsilon_i \stackrel{i.i.d.}{\sim} N(0, 0.01), x_i = \frac{i}{100}, i = 1, 2, \dots, 100.$$

考慮核修勻及 Spline 修勻，參數的選擇以上一題較佳者為原則，重複模擬 1000 次，計算不同方法下修勻值的偏誤(Bias)及變異數(Variance)，以此為標準判斷修勻方法的優劣。

3. 除了死亡率修勻，片段多項式也可用於探索資料特性。
 - (a) 仿造上課投影片，以 lowess 繪製臺灣 1960 年至今每年男女性平均壽命的延長趨勢。
 - (b) 請至 HMD (Human Mortality Database) 蒐集至少三個國家歷年平均壽命，仿造(a)的格式比較各國壽命延長趨勢。
4. 本題探討如何使用貝氏 (Kimeldorf-Jones) 修勻方法。
 - (a) 以修勻 2023 年簡易生命表為目標，請嘗試不同先驗分配 (Prior Distribution)，像是套用十年前同一縣市簡易生命表的死亡率、或十年前全臺灣簡易生命表的死亡率，比較兩種先驗分配的差異。(註：請同學選擇較佳之貝氏修勻參數。)
 - (b) 重複前一小題的修勻步驟，但人口數減少至原先之 1/20。
 - (c) 根據(a)的研究結果，以貝氏修勻編製 2019~2021 年國民生命表。