

110學年度高等水文分析作業一

第二部分—常態分佈與變數轉換(10分)

個人作業，請於11月8日(星期一) 3am以前將電子檔繳交至CEIBA網站
請第四、五組準備簡報11月11日8:10抽籤由一位同學第一節課報告作業成果

3. 利用第1題計算各年七月每天的日最高溫數據，以序號總和測試分析台北站1961-1980年、1990-2009年兩組20年的日最高溫紀錄，在 $\alpha=0.05$ 顯著水準條件下，判斷兩群日最高溫紀錄分布是否相同？（2分）
4. 在樣本資料實際上是呈極端值第一型分布，但虛無假設為「 H_0 ：隨機樣本呈常態分布」的「假設錯誤」類別問題中，選擇顯著水準 $\alpha=0.05$ ，利用蒙地卡羅法（根據多次重覆試驗結果，估計某種事件出現機率的方法），分析在不同樣本數條件下，分別找出卡方檢定發生第二型錯誤的機率，及其隨樣本數變化的曲線。設計卡方檢定時，請將常態分布分為8個等發生機率的區間，即每個區間內的發生機率都是0.125。（4分）
 - A. 利用uniform.for或其他程式，產生10,000組、每組 $8 \times n$ 個數值範圍為(0,1)的均勻分布隨機變數， $n=2,3,4,\dots$ ；將這10,000組、 $8 \times n$ 個的均勻分布隨機變數樣本值，視為是累積機率，轉換計算10,000組、 $8 \times n$ 個極端值第一型分布樣本。
 - B. 將每組 $8 \times n$ 個極端值第一型分布樣本，代入以上的卡方檢定，在虛無假設為 H_0 ：隨機樣本呈常態分布、顯著水準 $\alpha=0.05$ 條件下，得到每組數據接受或拒絕 H_0 的判斷。再綜合這10,000組資料的判斷結果，決定發生第二型錯誤的機率（即利用蒙地卡羅試驗得到的樣本機率）。持續增加 n 的數值， $n=2,3,4,\dots$ ，直到連續3個 n 值的10,000組的樣本第二型錯誤機率 β_s 都小於0.05時，便停止。
 - C. 以樣本數 $8 \times n$ 為橫軸（如果用Grapher繪圖，可以控制tic-label為8的倍數），蒙地卡羅試驗得到的樣本第二型錯誤機率為縱軸（兩個軸都可以採用線性或對數座標，重點是要能夠最清楚呈現想要觀察的結果，即曲線變化和曲線通過0.05的情形），繪圖表現蒙地卡羅試驗的結果。
5. 二維卡氏座標位置分別在O(0,0)、A(20,10)、B(25, 50)、C(-80、30)、

D(-20, -60)的五個溫度測站，座標數值的單位都是公里。各測站的日均溫均為常態分佈的隨機變數，期望值均為30度、標準偏差均為3度，相關係數是距離的函數 $\rho(d) = \exp(-d/30)$ 。(4分)

A. 某日O站和D站的日均溫分別為33.3度和29.7度，請「最佳化估計」

A、B、C三站的溫度，計算三個估計值不確定性的變異數，以及 $\rho_{AB|OD}$ 、 $\rho_{AC|OD}$ 、 $\rho_{BC|OD}$ 分別是否與 ρ_{AB} 、 ρ_{AC} 、 ρ_{BC} 相同？為何？

B. 若某日O站缺測，擬用 $\hat{T}_O = \sum_i w_i T_i$ ， $i = A、B、C、D$ 補遺估計，請決定四個測站的權重係數值 w_i 。