**國立台北護理健康大學**

**流行病學方法論及實驗：期末測驗練習**

**學號 　　　　　 姓名　　　　　　　　 2019-06-14**

有一研究欲了解飲用含糖飲料之習慣對於高血壓之影響。該研究共納入台灣某社區民眾2,036人，所蒐集之資料(檔名：data.sav)包含年齡、性別(gender, 1:男性/2:女性)、是否飲用含糖飲料(drk, 1:有/0:無)、是否有高血壓及是否具抽菸習慣(smk, 1:有或曾經/0:無)。請利用此資料回答以下問題：

1. **第一部分：疾病盛行率及發生率**
2. 試問飲用含糖飲料與未具有此習慣者的高血壓盛行率(Prevalence)為何？
3. 飲用含糖飲料：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. 不飲用含糖飲料：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
5. 假設高血壓發生率服從一指數分佈且年發生率()為0.12，試回答以下問題：
6. 試計算並解釋高血壓PI ratio？\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
7. 試計算高血壓平均病程時間？\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
8. 試問多少年後會多少一半無高血壓族群發生高血壓？\_\_\_\_\_\_\_\_\_\_\_\_\_\_
9. 試繪製無病存活曲線(可利用0308 Excel)
10. **第二部分：辛普森矛盾**
11. 試問含糖飲料飲用分佈(比例)在高血壓及無高血壓族群是否有所不同？
12. 檢定方法：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
13. 統計量及P值：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
14. 解釋：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
15. 試利用「計算變數」將族群以50歲分為兩個年齡層，令變項名稱為Agegroup，其中年輕組為age<50，年老組為Age>=50。
16. 試利用分層分析(年輕組及年老組)，評估含糖飲料飲用分佈(比例)在高血壓及無高血壓族群是否有所不同？
17. 檢定方法：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
18. 統計量及P值：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
19. 解釋：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(提示：是否有交互作用)
20. 飲用含糖飲料比上不飲用含糖飲料的高血壓勝算比為何？

年輕組：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

年老組：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

合併：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (若無法進行合併請打X)

1. 依照2.欲檢定內容，建構兩個羅吉斯迴歸模式：
2. 模式1：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. 模式2：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. 模式之 -2對數概似值 (-2LL)：

* 模式1：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* 模式2：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. 依照您建構之模型寫下欲檢定之參數以及假說：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. 兩模型概似值檢定對於(4)之結果為何?
3. 寫下飲用含糖飲料對於高血壓之影響之估計結果及其95%信賴區間為何。

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. 傾向分數分析
2. 試利用年齡、性別及抽菸習慣建立一飲用含糖飲料之傾向分數。
3. 試將此傾向分數放入迴歸模式中進行調整，寫下飲用含糖飲料對於高血壓之影響之估計結果及其95%信賴區間為何。

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. 利用此傾向分數進行配對(檔名：data\_ps.sav)。試利用此資料集，寫下飲用含糖飲料對於高血壓之影響之估計結果及其95%信賴區間為何。

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **第三部分：科學臨床推理**

利用年齡、性別、是否飲用含糖飲料及是否抽菸建構一多變項羅吉斯迴歸，試回答以下問題：

1. 若有一民眾，其特性如下：56歲，女性，具有飲用含糖飲料之習慣；具抽菸習慣)。 利用模型估計結果，計算該民眾罹患高血壓之機率。\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. 試建立此預測模型之ROC曲線並寫下其曲線下面積(AUC)。
3. 實際運用此預測模型時，若將高血壓切點值訂於50%(>=50%)為陽性，試完成下表：

|  |  |  |
| --- | --- | --- |
|  | 實際高血壓 | 實際無高血壓 |
| 模型預測高血壓 |  |  |
| 模型預測無高血壓 |  |  |

1. 試問此模型對於預測高血壓之敏感度(sensitivity)、特異度(specificity)、陽性預測值(PPV)、陰性預測值(NPV)分別為何? 敏感度/(1-特異度)為何?
2. 依據過去資料，曾有30%的民眾有高血壓情形。
3. 試問高血壓勝算為何(事前勝算)?
4. 若有一位民眾，運用此預測模式且結果呈現具有高血壓(預測機率>50%)，基於此一結果，則該民眾患有高血壓的勝算與機率分別為何?