**統計諮詢\_作業七**

統計所碩一 R26091024王苡璿

**Exercises 15.4**

一、資料說明

資料集(mortality)提供了1974年紐約市37,840名非白人母親分娩之新生兒存活概況，並以出生後一年以及出生重量2500克作為分類基準。

欲建構樣本勝算比(odds ratio)及解釋其代表意義，並建構母體勝算比之95%信賴區間，資料列聯表如表1所示，由此列聯表可以得到在存活超過一年的嬰兒中，大於2500克的嬰兒佔了88.26%；而存活不到一年的嬰兒中，小於2500克的嬰兒佔了38.59%。

**表 1：資料列聯表**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 死亡 | 存活 | 總和 |
| 小於等於2500克 | 530 | 4340 | 4870 |
| 大於2500克 | 333 | 32637 | 32970 |
| 總和 | 863 | 36977 | 37840 |

二、勝算比(odds ratio)

1. 定義

1. 勝算(odds)

發生某事件比率與未發生該事件比率之比值，以此資料集為例，小於等於2500克的嬰兒死亡之勝算(odds)為0.1221。

1. 勝算比(odds ratio)

兩個勝算(odds)之比值即為勝算比(odd ratio)，以此資料集為例，即是小於等於2500克的嬰兒死亡之勝算，對大於2500克的嬰兒死亡之勝算比值。

2. 樣本勝算比(sample odds ratio)

由表2可得，出生未達一年死亡的嬰兒出生體重小於等於2500克的勝算

是大於2500克的11.96倍，如果在更多的樣本數下，仍粗略維持此勝算

比，則可以推測可能小於等於2500克的嬰兒較大於2500克的嬰兒容易

壽命維持不到一年。

**表 2：樣本勝算與勝算比**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 死亡 | 存活 | 嬰兒死亡之勝算(odds) |
| 小於等於2500克 | 530 | 4340 | 0.1221 |
| 大於2500克 | 333 | 32637 | 0.0102 |
| 勝算比(odds ratio) | - | - | 11.96 |

三、勝算比(odds ratio)之95%信賴區間

1. 在大樣本下之漸進分配

令代表母體勝算比，代表估計母體勝算比之估計量，則在大樣本下，

ln() ~ N( ln() , )，其中，、、 、代表資料列聯表中四個數據資料。

2. 母體勝算比之95%信賴區間

ln()之95%信賴區間如下

CI( ln() ) = ln() = (L , U)，

而之95%信賴區間即為CI( ) = ()。

由以上之定義可計算出此資料集之母體勝算比95%信賴區間為

(10.398,13.777)。

**Exercises 17.4**

一、資料說明

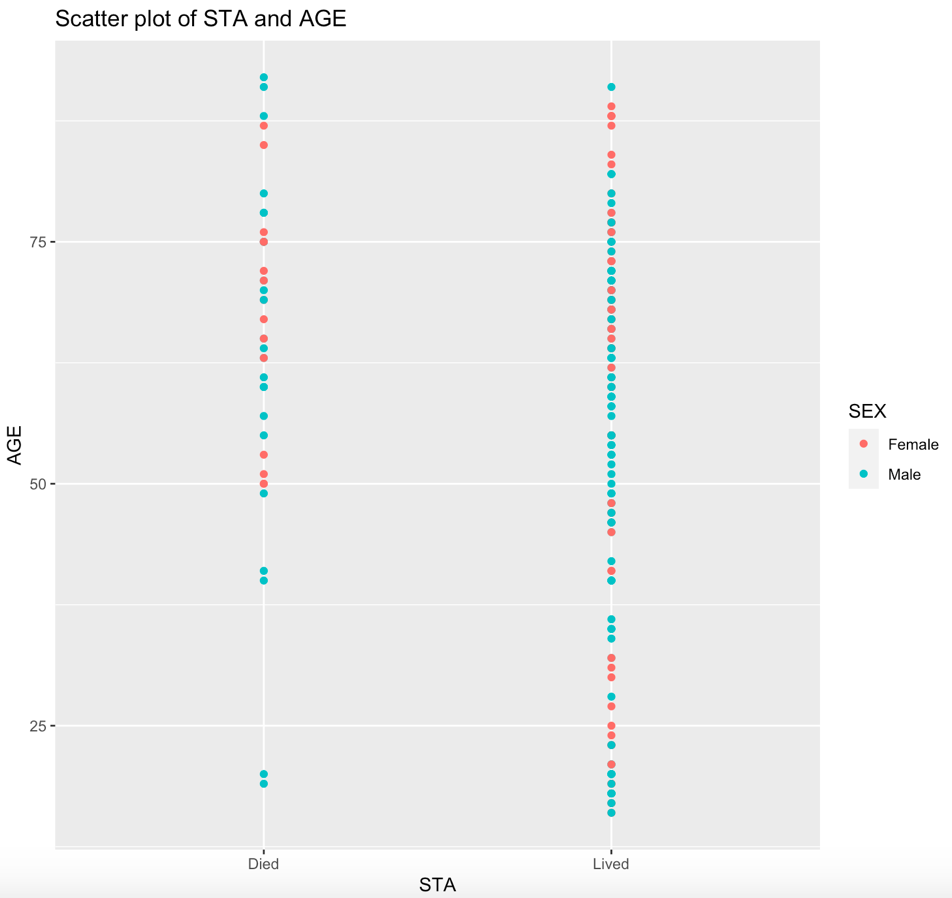
資料集(icu)是蒐集加護病房中病患的資訊，其中包含病患編號、生命狀態、性別、年齡、人種等等，總共21項資訊。由於諮詢者特別對性別及年齡對於生命狀態之關係感興趣，因此僅針對此三個變數做介紹，詳細變數介紹如表1所示。

**表 1：變數介紹**

|  |  |  |
| --- | --- | --- |
| 變數  名稱 | 意義說明 | 變數  類型 |
| STA | 生命狀態  (Lived：存活、Died：死亡) | 屬質  變數 |
| AGE | 年齡 | 屬量  變數 |
| SEX | 性別  (Male：男性、Female：女性) | 屬質  變數 |

由圖1可得，在死亡的病患中，低於五十歲以下的僅有少數男性，沒有女性病患；在存活的病患中，男性較女性多。

**圖 1：生命狀態與年齡在性別條件下散佈圖**

****

二、羅吉斯迴歸(Logistic regression)

1.模型

，

其中代表第個觀察值生命狀態為死亡之機率，代表第個觀察值之性

別，代表第個觀察值之年齡。

2.模型配適結果

= 由模型結果可得到：

1. 在其他條件固定下，當病患為女性時，使得死亡的勝算降低約0.01倍()
2. 在其他條件固定下，當病患年齡增加一歲，使得死亡的勝算提高約1.03倍()

**表 2：羅吉斯迴歸模型配適結果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 羅吉斯迴歸模型配適 | | | | |
|  | 估計值 | 標準誤 | 檢定統計量 | p-value |
|  | -3.05669 | 0.69884 | -4.374 | 1.22e-05 |
|  | -0.01131 | 0.37177 | -0.030 | 0.97573 |
|  | 0.02758 | 0.01066 | 2.589 | 0.00963 |

三、模型比較

1.模型建立與配適結果

模型I（不考慮性別與年齡之交互作用）

，

模型配適結果：=

模型II（考慮性別與年齡之交互作用）

，

模型配適結果：=

結論：考慮性別與年齡交互作用項後，在其他條件固定下，當病患為女

性時，死亡的勝算從降低轉變為提升；在其他條件固定下，性別

與年齡之交互作用，使得死亡的勝算降低。

2.利用概似比例檢定(likelihood ratio test)比較兩個模型

：模型I之擬合程度較模型II佳

：模型II之擬合程度較模型II佳

由表3之檢定結果可得，沒有顯著證據顯示模型I之擬合程度較模型II

佳，因此我們將選用模型II之配適結果。

**表 3：概似比例檢定比較模型結果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Analysis of Deviance Table | | | | | |
| Model 1：STA ~ SEX + AGE  Model 2：STA ~ SEX \* AGE | | | | | |
|  | Resid. Df | Resid. Dev | Df | Deviance | Pr(>Chi) |
| Model 1 | 197 | 192.31 |  |  |  |
| Model 2 | 196 | 192.30 | 1 | 0.0012861 | 0.9714 > 0.05 |