* 流病學原理資料分析課堂作業3

流行病學資料分析 – SAS

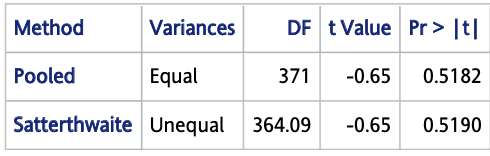
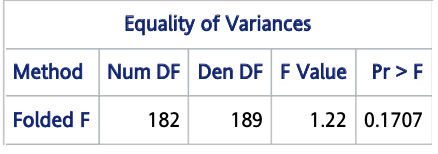
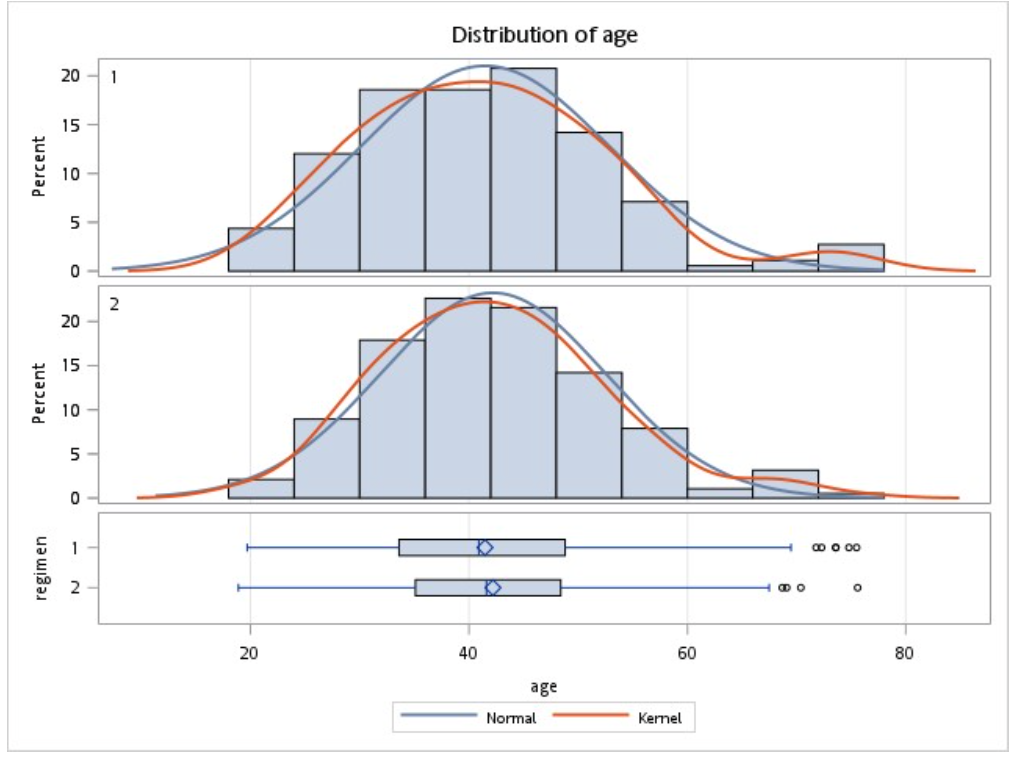
Practice 3

學號：b07401048

系級：醫學五

姓名：賴柏瑞

* 1. 請比較被隨機分派到INH與RMP的兩組LTBI病人？
     1. 年齡
        1. Results：
           1. 由於年齡為等距變項，故以兩獨立樣本t檢定判斷INH與RMP兩組織年齡是否有顯著差異。
           2. 先以folded F判斷兩者之變異數是否相等（如下左圖）。虛無假說為兩者變異數沒有顯著差異，對立假設為兩者變異數有顯著差異。檢定結果之p-value為0.1707，無法拒絕虛無假說，兩組變異數沒有顯著差異。
           3. 之後以假定兩組變異數相等的pooled two-sampled t-test檢定兩組年齡平均數是否有顯著差異（如下右圖）。虛無假說為兩者沒有年齡平均數沒有顯著差異，對立假說為兩者年齡平均數有顯著差異。檢定結果之p-value為0.5182，無法拒絕虛無假說，兩組年齡平均數沒有顯著差異。
           4. 年齡為影響肝毒性風險的重要因素，在隨機分派的試驗中，是必須需確保實驗組與對照組的年齡相等，以免影響結果判讀。**INH與RMP兩組的年齡平均數沒有顯著差異，**符合隨機分派的標準。
        2. Figures：



* + - 1. Code：

*/\* homework 3 \*/*

*/\* Q1 \*/*

*/\* Q1-1 : AGE <- t-test \*/*

title Q1\_1;

**proc** **ttest** data = dataset1;

class regimen;

var age;

**run**;

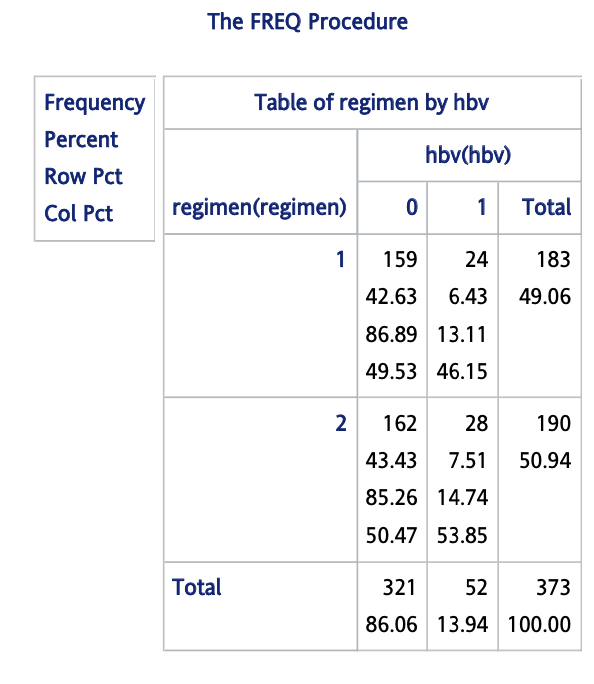
*/\* extra: check if age is in normal distribution \*/*

**proc** **univariate** data = dataset1 normal plot;

var age;

**run**; */\* actually not in norm, but we have big data (300 > 30), can ignore*

* + 1. HBV
       1. Result：
          1. HBV帶原（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組之有無HBV帶原是否有顯著差異
          2. 由左圖之列聯表可見INH組（regimen 1）HBV帶原的比例為13.11%，RMP組（regimen 2）HBV帶原的比例為14.74%。
          3. 以卡方檢定判斷兩組HBV帶原的比例是否有顯著差異，虛無假說為HBV帶原與分組為獨立事件，對立假說為HBV帶原與分組有相關性。檢定結果之p-value為0.6512，無法拒絕虛無假說，表示HBV帶原與分組為獨立事件，兩組HBV帶原的比例沒有顯著差異。
          4. HBV帶原會導致慢性肝炎，與藥物肝毒性的感受性有重大關係，在隨機分派的試驗中需確保兩組的HBV帶原率相等，以免影響結果判讀。**INH與RMP兩組的HBV帶原率沒有顯著差異**，符合隨機分派的標準。
       2. Figures：



* + - 1. Code：

*/\* Q1-2 : HBV <- chi-square \*/*

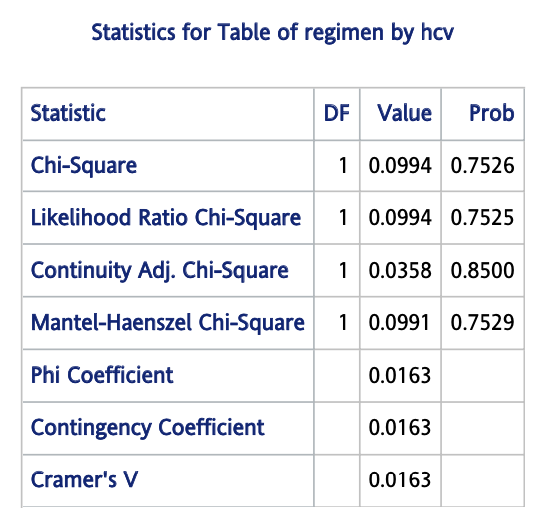
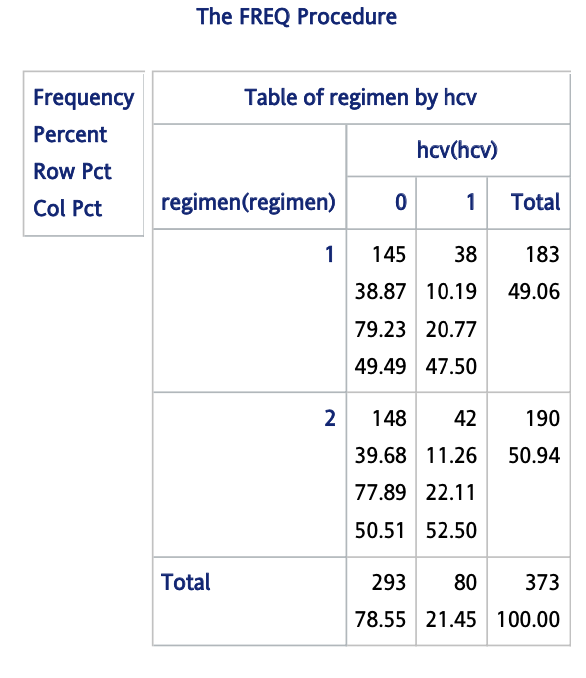
title Q1\_2;

**proc** **freq** data = dataset1;

tables regimen \* hbv / chisq;

**run**;

* + 1. HCV：
       1. Result：
          1. HCV帶原（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組之有無HCV帶原率是否有顯著差異
          2. 由左圖之列聯表可見INH組（regimen 1）HCV帶原的比例為20.77%，RMP組（regimen 2）HCV帶原的比例為22.11%。
          3. 以卡方檢定判斷兩組HCV帶原的比例是否有顯著差異，虛無假說為HCV帶原與分組為獨立事件，對立假說為HCV帶原與分組有相關性。檢定結果之p-value為0.7526，無法拒絕虛無假說，表示HCV帶原與分組為獨立事件，兩組HBV帶原的比例沒有顯著差異。
          4. HCV會導致慢性肝炎，與藥物肝毒性的感受性有重大關係，在隨機分派的試驗中需確保兩組的HCV帶原率相等，以免影響結果判讀。INH與RMP兩組的HCV帶原率沒有顯著差異，符合隨機分派的標準。
       2. Figures：



* + - 1. Code：

*/\* Q1-3 : \*/*

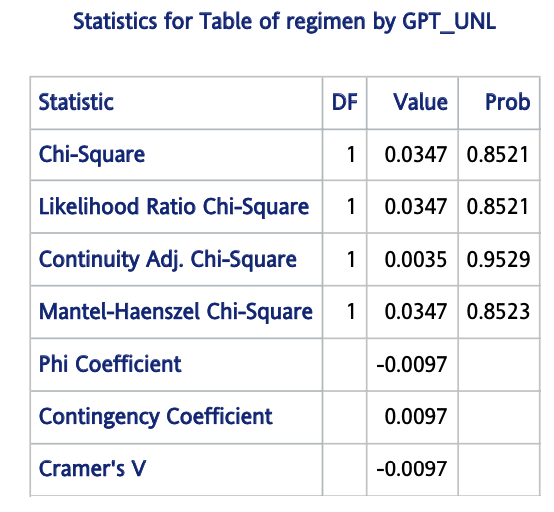
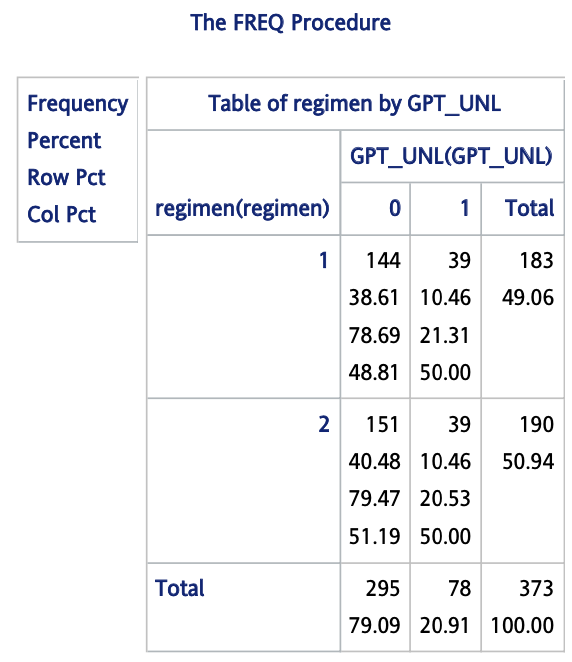
title Q1\_3;

**proc** **freq** data = dataset1;

tables regimen \* hcv / chisq;

**run**;

* + 1. 收案時的肝指數是否有異常
       1. Result：
          1. 收案時肝指數（GPT）是否有異常（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組病人在收案時肝指數有異常的比例是否有顯著差異。
          2. 由左圖之列聯表可見INH組（regimen 1）收案時肝指數有異常的比例為21.31%，RMP組（regimen 2）收案時肝指數有異常的比例為20.53%。。
          3. 以卡方檢定判斷兩組收案時肝指數有異常的比例是否有顯著差異，虛無假說為收案時肝指數有異常與分組為獨立事件，對立假說為收案時肝指數有異常與分組有相關性。檢定結果之p-value為0.8521，無法拒絕虛無假說，表示收案時的肝指數異常與分組為獨立事件，兩組收案時肝指數異常的比例沒有顯著差異。
          4. 肝指數可代表baseline 的肝功能，與藥物肝毒性的感受性有重大關係，在隨機分派的試驗中需確保兩組baseline肝功能相等，以免影響結果判讀。**INH與RMP兩組收案時肝指數異常率沒有顯著差異**，符合隨機分派的標準。
       2. Figures：



* + - 1. Code：

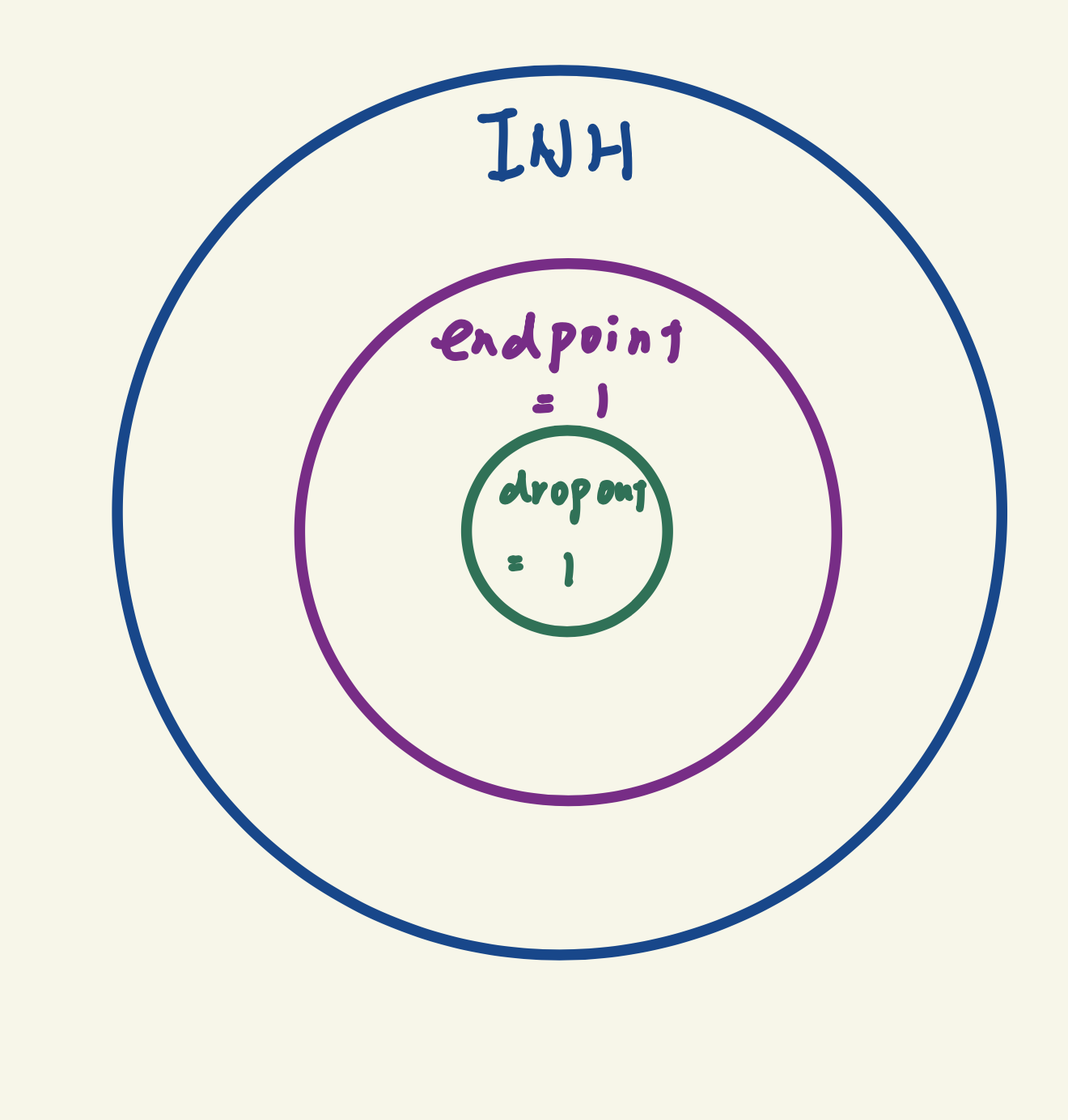
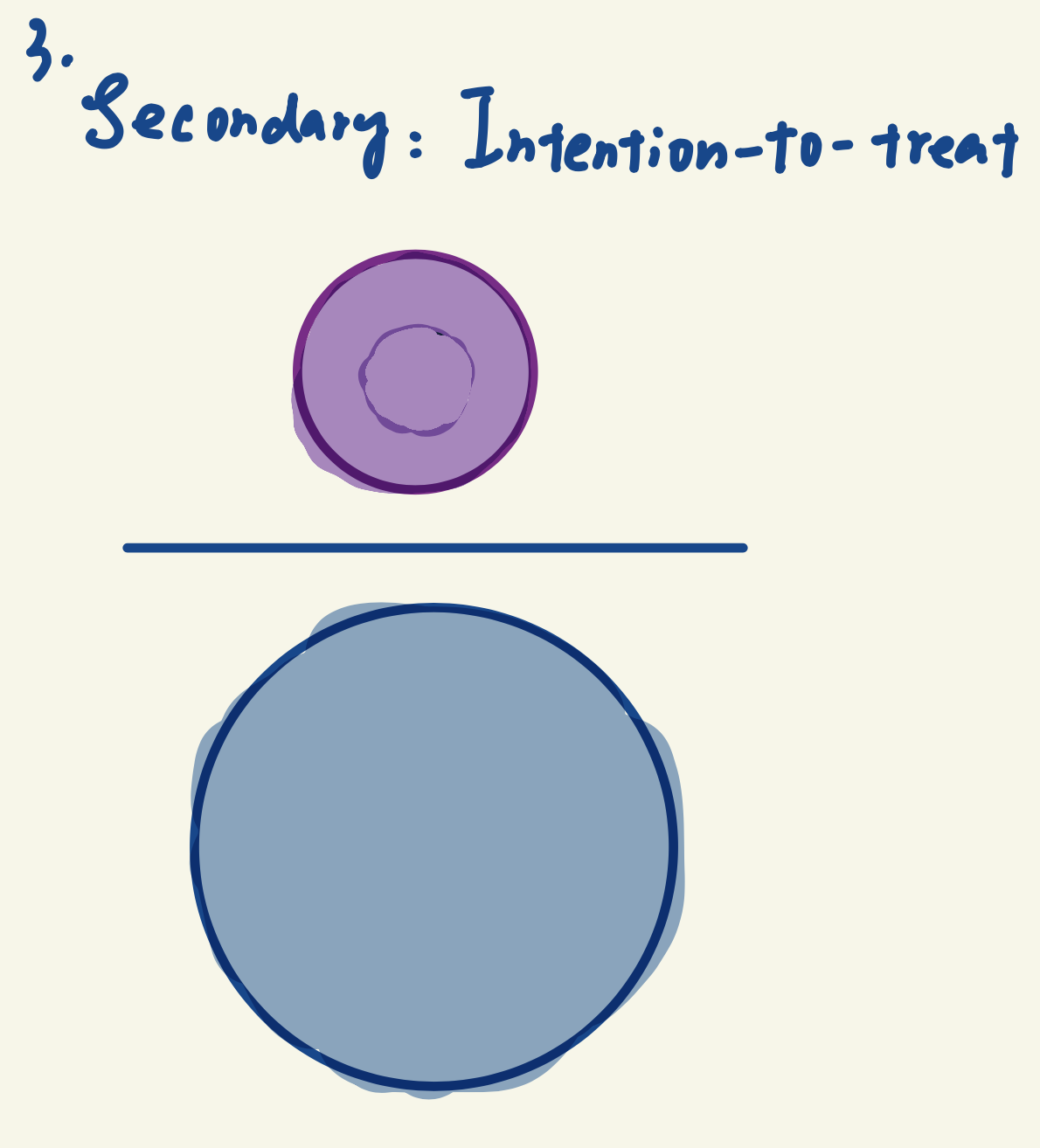
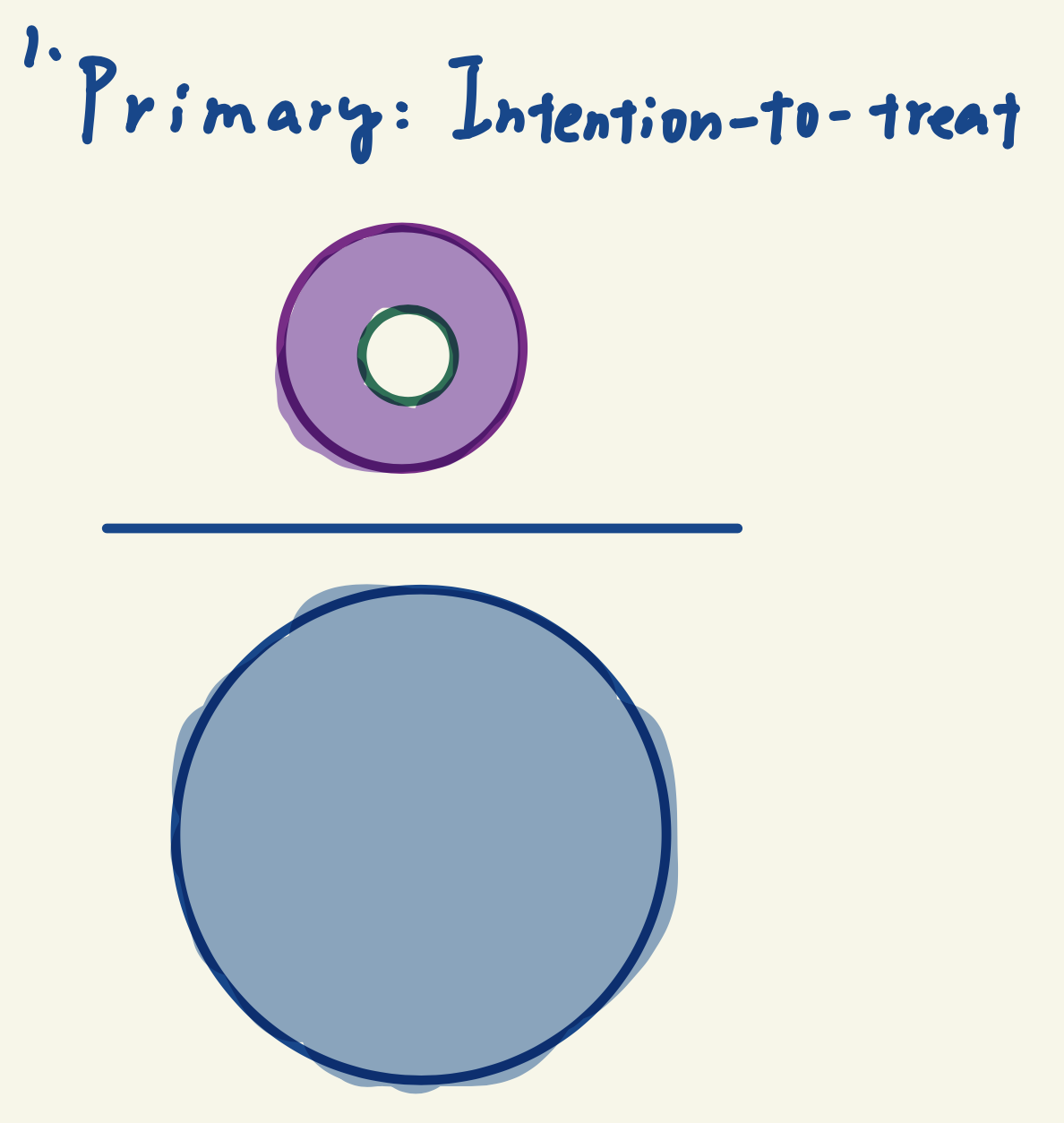
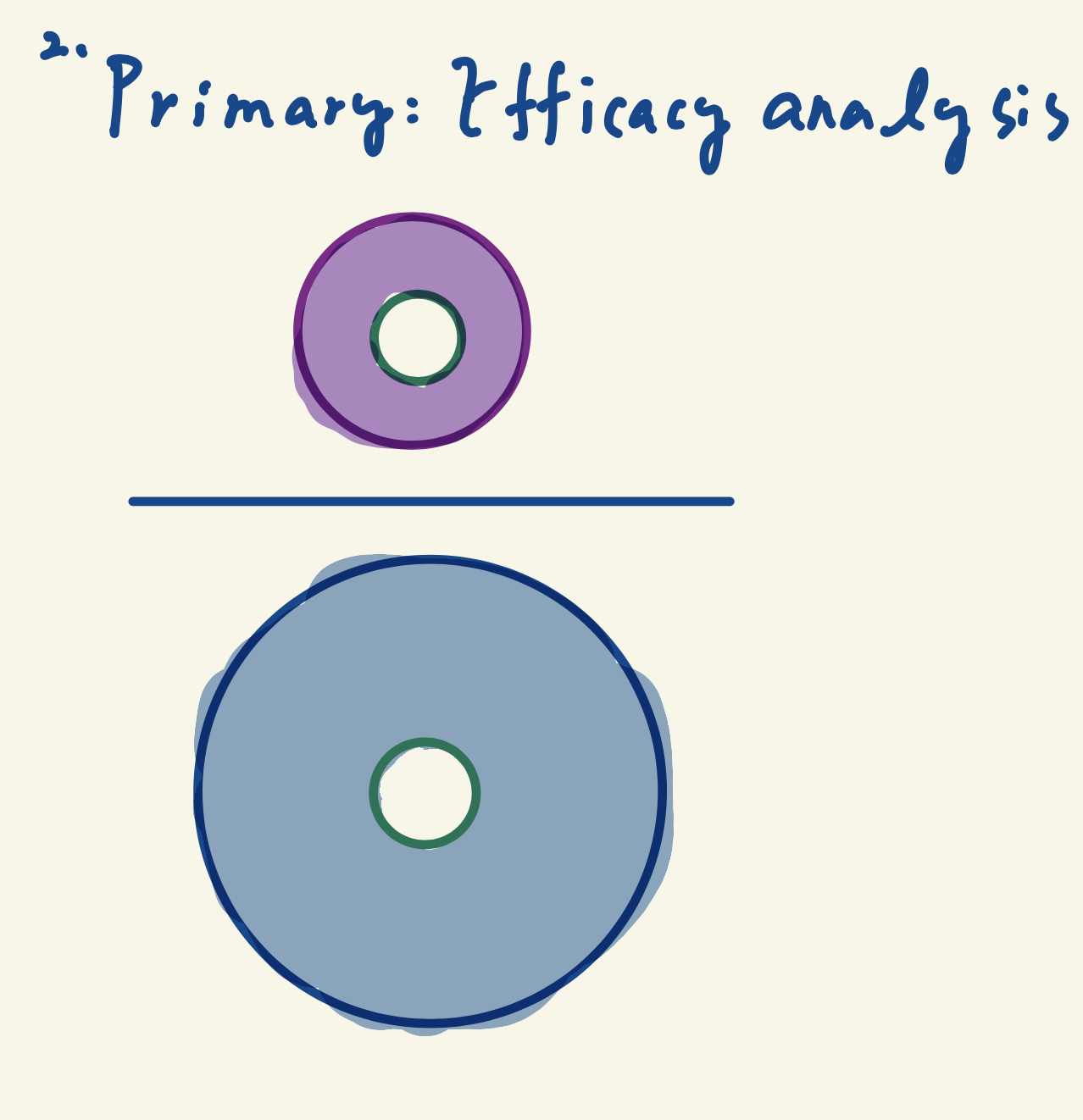
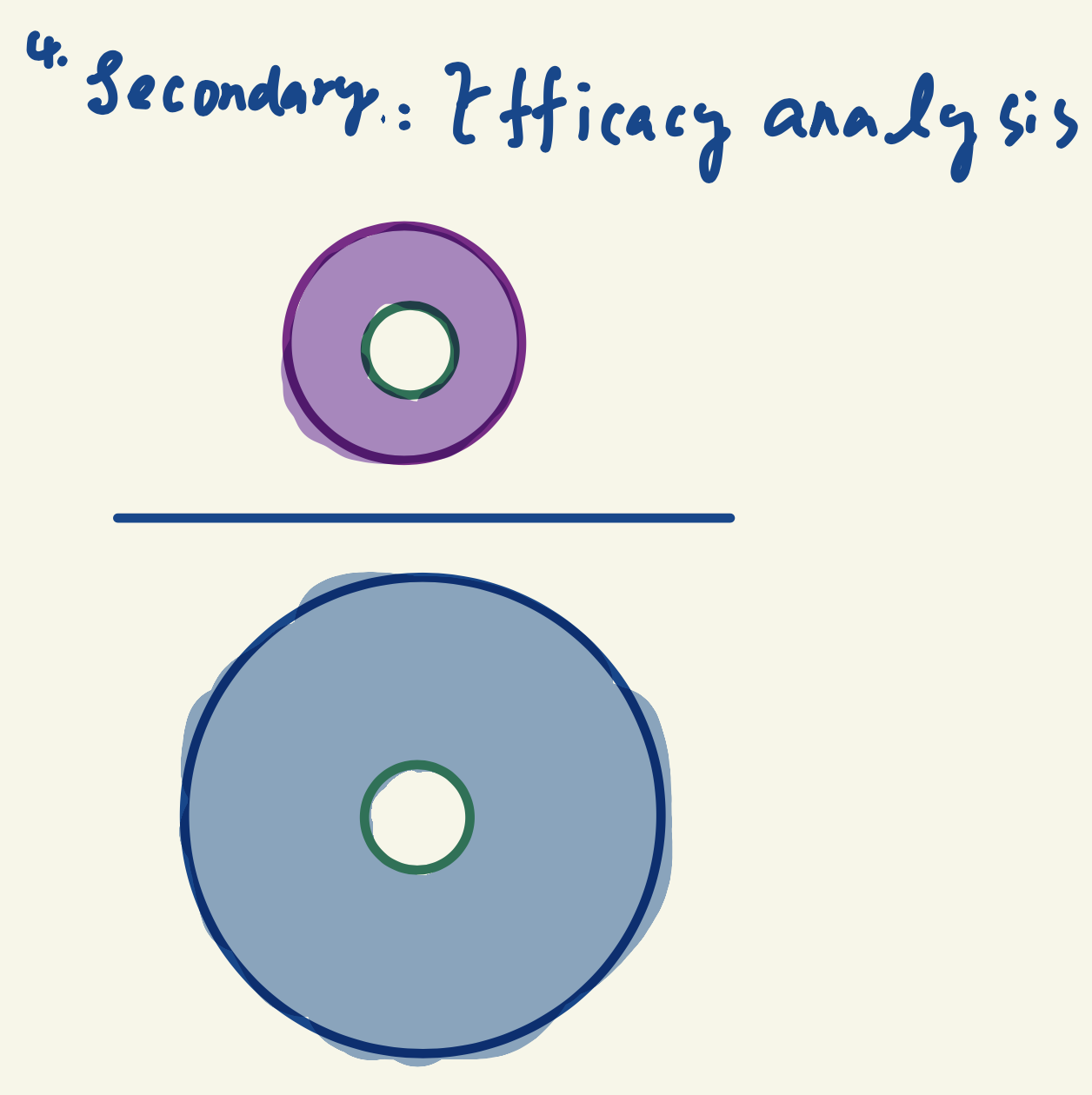
*/\* Q1-4 : \*/*

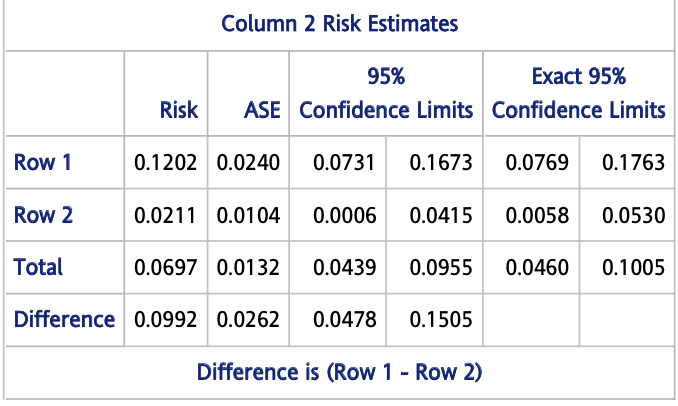
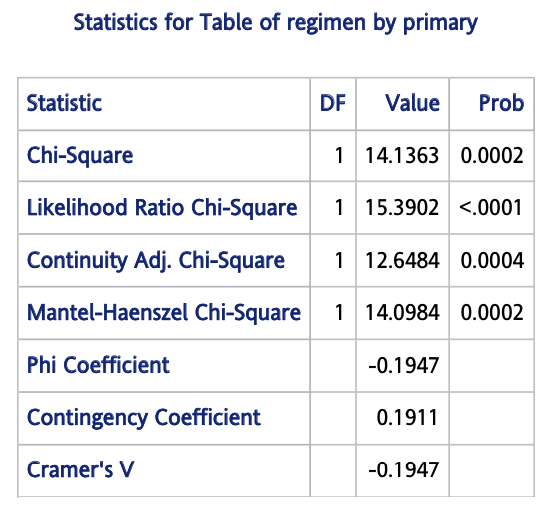
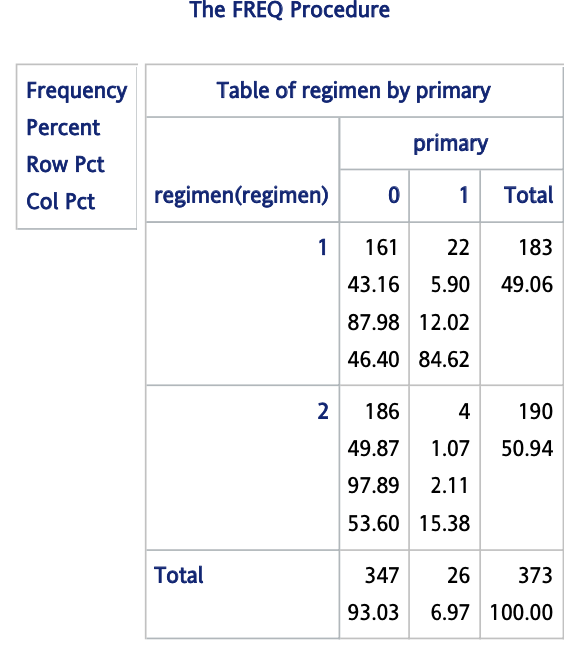
title Q1\_4;

**proc** **freq** data = dataset1;

tables regimen \* GPT\_UNL / chisq;

**run**;

* 1. 請比較以下評估項目RMP與INH兩組是否有統計上的顯著差異？
     1. 簡圖：
     2. Primary endpoint, by intention-to-treat
        1. Result：
           1. 本試驗的primary endpoint定義為「因藥物肝毒性或其他嚴重副作用而被迫停藥」，intention-to-treat則表示以收案時該組人數表示該組的總人數。
           2. 是否達到primary endpoint（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組病人達到primary endpoint的比例是否有顯著差異。
           3. 由左圖之列聯表可見INH組（regimen 1）達到primary endpoint的比例為12.02%，RMP組（regimen 2）達到primary endpoint的比例為2.11%。
           4. 同時，也可以risk敘述：INH組病人在INH療程的時間內，發生primary endpoint的risk為12.02%；RMP組在RMP療程的時間內，發生primary endpoint的risk為2.11%，兩組的risk difference為9.92%。
           5. 以卡方檢定判斷兩組達到primary endpoint的比例是否有顯著差異，虛無假說為達到primary endpoint與分組為獨立事件，對立假說達到primary endpoint與分組有相關性。檢定結果之p-value為0.0002，拒絕虛無假說，**表示兩組達到primary endpoint的比例有顯著差異**。
        2. Figures：



* + - 1. Code：

*/\* Q2 \*/*

*/\* Q2-1 primary endpoint by intention-to-treat\*/*

title Q2\_1;

**data** primary;

set dataset1;

if endpoint - dropout = 1 then primary = 1;

else primary = 0;

**run**;

**proc** **print** data = primary;

**run**;

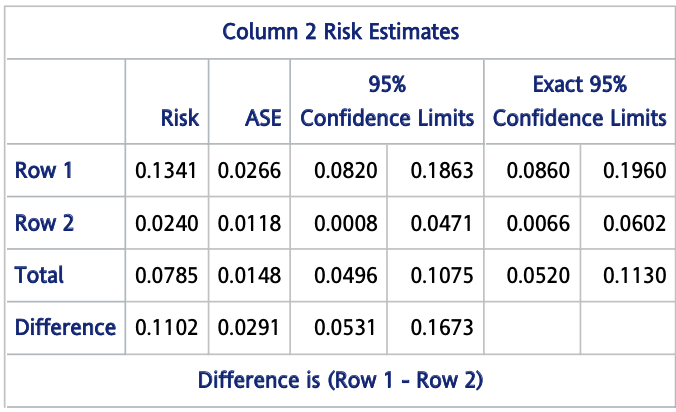
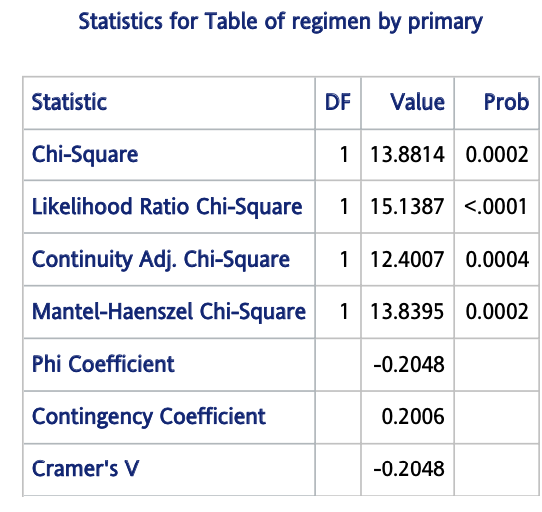
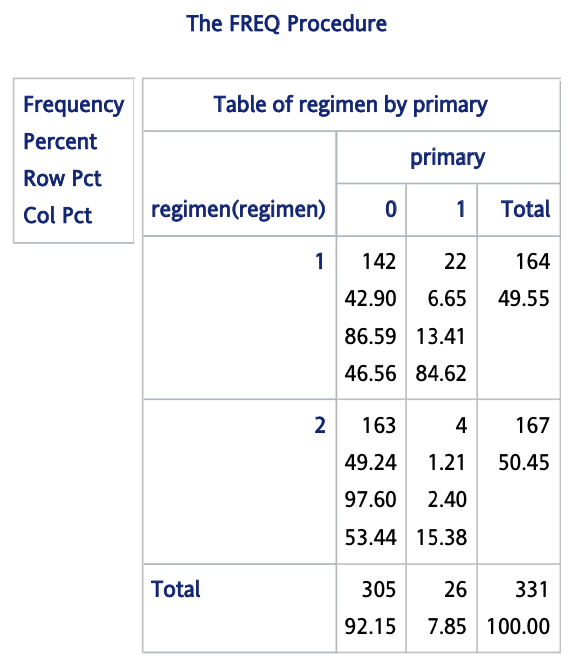
*/\* chisqure + risk difference \*/*

**proc** **freq** data = primary;

tables regimen \* primary / chisq riskdiff;

**run**;

* + 1. Primary endpoint, by efficacy analysis
       1. Result：
          1. 本試驗的primary endpoint定義為「因藥物肝毒性或其他嚴重副作用而被迫停藥」，efficacy則表示不納入中途dropout的人，以完成療程的病人為該組的人數。
          2. 是否達到primary endpoint（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組病人達到primary endpoint的比例是否有顯著差異。
          3. 由左圖之列聯表可見INH組完成療程的病人中（regimen 1）達到primary endpoint的比例為13.41%，RMP組完成療程的病人中（regimen 2）達到primary endpoint的比例為2.40%。
          4. 同時，也可以risk敘述：INH組病人在完成INH療程的時間內，發生primary endpoint的risk為13.41%；RMP組在完成RMP療程的時間內，發生primary endpoint的risk為2.40%，兩組的risk difference為11.02%。
          5. 以卡方檢定判斷兩組達到primary endpoint的比例是否有顯著差異，虛無假說為達到primary endpoint與分組為獨立事件，對立假說達到primary endpoint與分組有相關性。檢定結果之p-value為0.0002，拒絕虛無假說，**表示兩組達到primary endpoint的比例有顯著差異**。
       2. Figures：



* + - 1. Code：

*/\* Q2-2 primary endpoint by efficacy analysis \*/*

title Q2\_2;

**data** primary\_efficacy;

set dataset1;

if dropout = 1 then delete;

if endpoint = 1 then primary = 1;

else primary = 0;

**run**;

**proc** **print** data = primary\_efficacy;

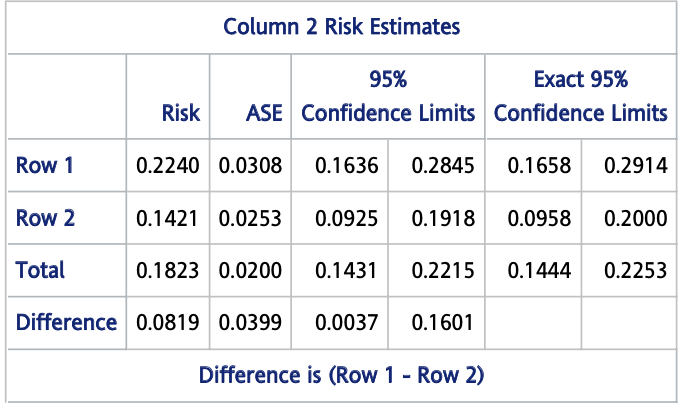
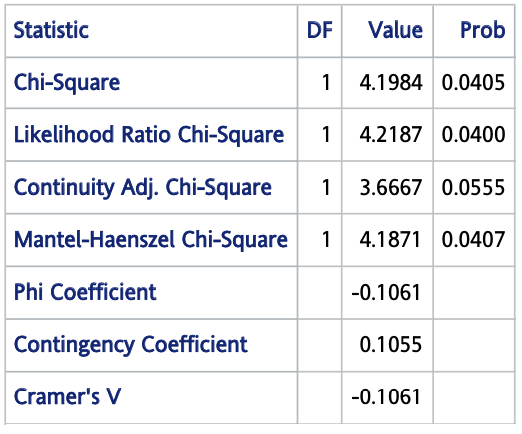
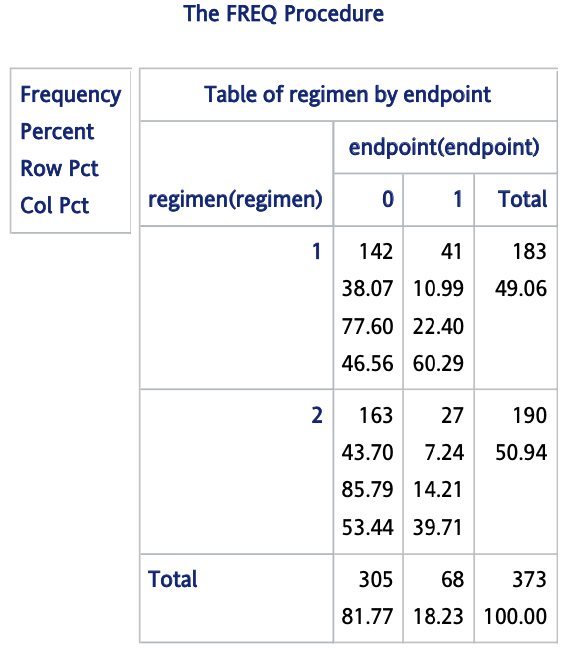
**run**;

**proc** **freq** data = primary\_efficacy;

tables regimen \* primary / chisq riskdiff;

**run**;

* + 1. Secondary endpoint, by intention-to-treat
       1. Result：
          1. 本試驗的secondary endpoint定義為「因任何原因而停藥（包含drop out）」，intention-to-treat則表示以收案時該組人數表示該組的總人數。
          2. 是否達到secondary endpoint（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組病人達到secondary endpoint的比例是否有顯著差異。
          3. 由左圖之列聯表可見INH組（regimen 1）達到secondary endpoint的比例為22.40%，RMP組（regimen 2）達到secondary endpoint的比例為14.21%。
          4. 同時，也可以risk敘述：INH組病人在INH療程的時間內，發生secondary endpoint的risk為22.40%；RMP組在RMP療程的時間內，發生primary endpoint的risk為14.21%，兩組的risk difference為8.19%。
          5. 以卡方檢定判斷兩組達到secondary endpoint的比例是否有顯著差異，虛無假說為達到secondary endpoint與分組為獨立事件，對立假說達到secondary endpoint與分組有相關性。檢定結果之p-value為0.0405，拒絕虛無假說，**表示兩組達到secondary endpoint的比例有顯著差異**。
       2. Figures：



* + - 1. code

*/\* Q2-3 secondary endpoint by intention-to-treat \*/*

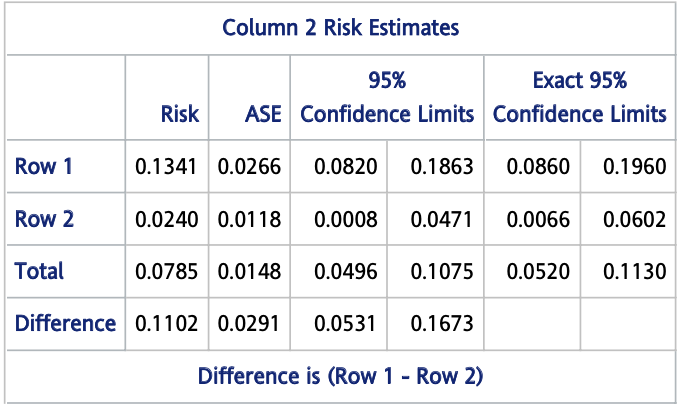
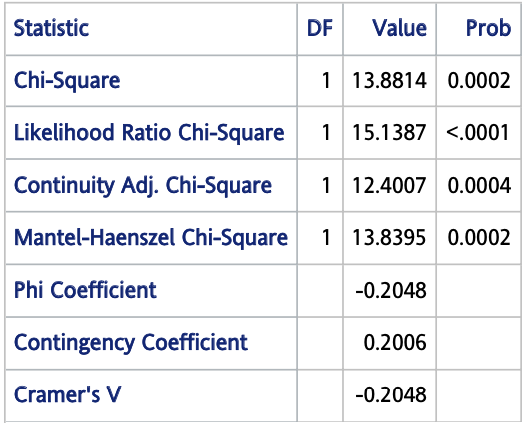
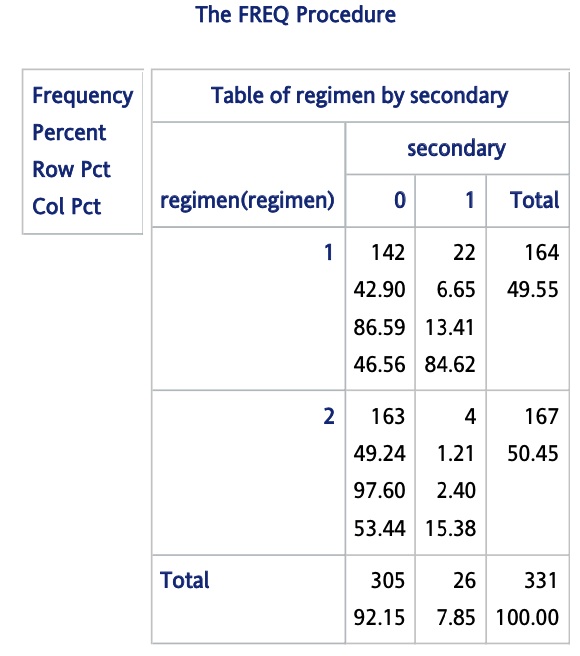
title Q2\_3;

**proc** **freq** data = dataset1;

tables regimen \* endpoint / chisq riskdiff;

**run**;

* + 1. Secondary endpoint, by efficacy analysis
       1. Result：
          1. 本試驗的secondary endpoint定義為「因任何原因而停藥」，efficacy則表示不納入中途dropout的人，以完成療程的病人為該組的人數。
          2. 是否達到secondary endpoint（有無）為類別變項，故以卡方檢定判斷RMP與INH兩組病人達到secondary endpoint的比例是否有顯著差異。
          3. 由左圖之列聯表可見INH組完成療程的病人中（regimen 1）達到secondary endpoint的比例為13.41%，RMP組完成療程的病人中（regimen 2）達到secondary endpoint的比例為2.40%。
          4. 同時，也可以risk敘述：INH組病人在完成INH療程的時間內，發生secondary endpoint的risk為13.41%；RMP組在完成RMP療程的時間內，發生secondary endpoint的risk為2.40%，兩組的risk difference為11.02%。
          5. 以卡方檢定判斷兩組收案時肝指數有異常的比例是否有顯著差異，虛無假說為達到secondary endpoint與分組為獨立事件，對立假說達到secondary endpoint與分組有相關性。檢定結果之p-value為0.0002，拒絕虛無假說，**表示兩組達到secondary endpoint的比例有顯著差異**。
       2. Figures：



* + - 1. Code：

*/\* Q2-4 secondary endpoint by efficacy analysis \*/*

title Q2\_4;

**data** secondary\_efficacy;

set dataset1;

if dropout = 1 then delete;

if endpoint = 1 then secondary = 1;

else secondary = 0;

**run**;

**proc** **print** data = secondary\_efficacy;

**run**;

**proc** **freq** data = secondary\_efficacy;

tables regimen \* secondary / chisq riskdiff;

**run**;