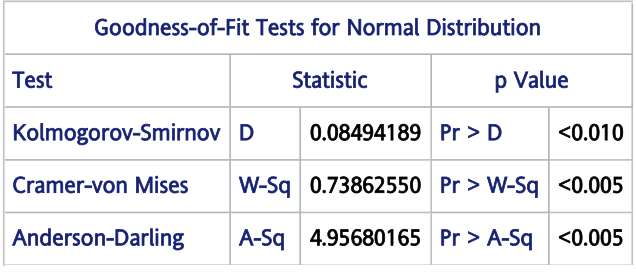
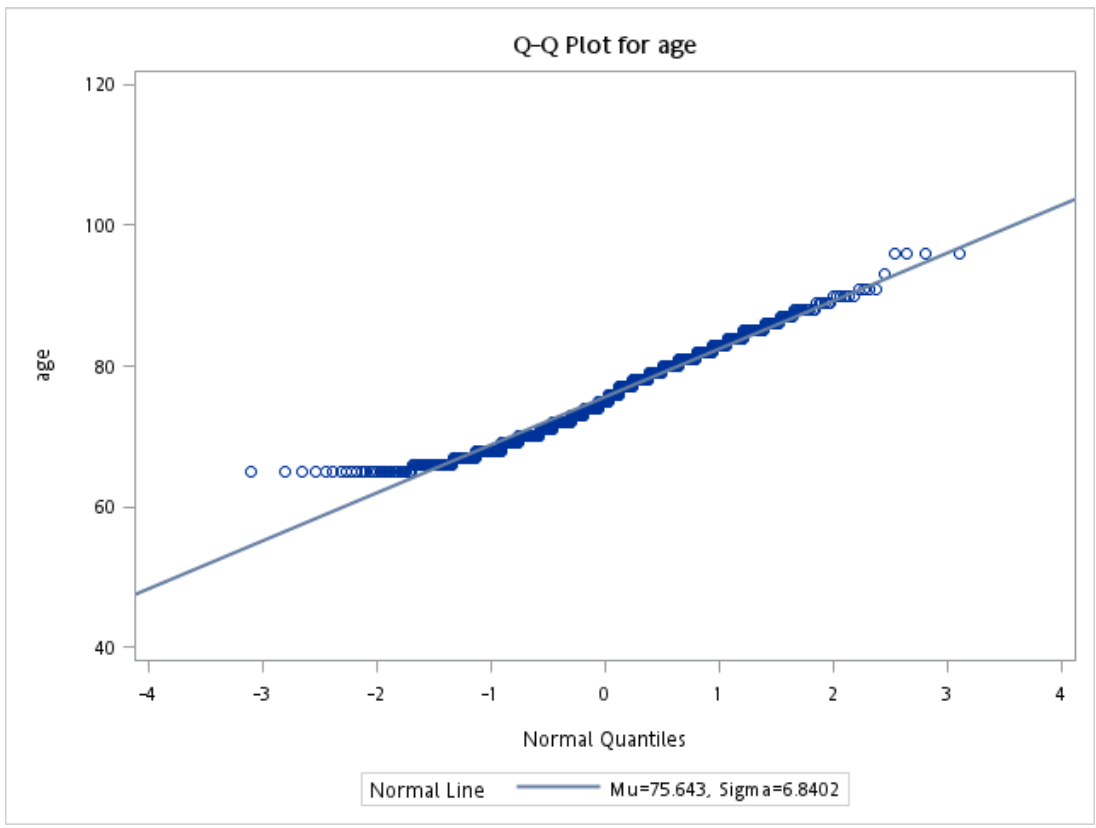
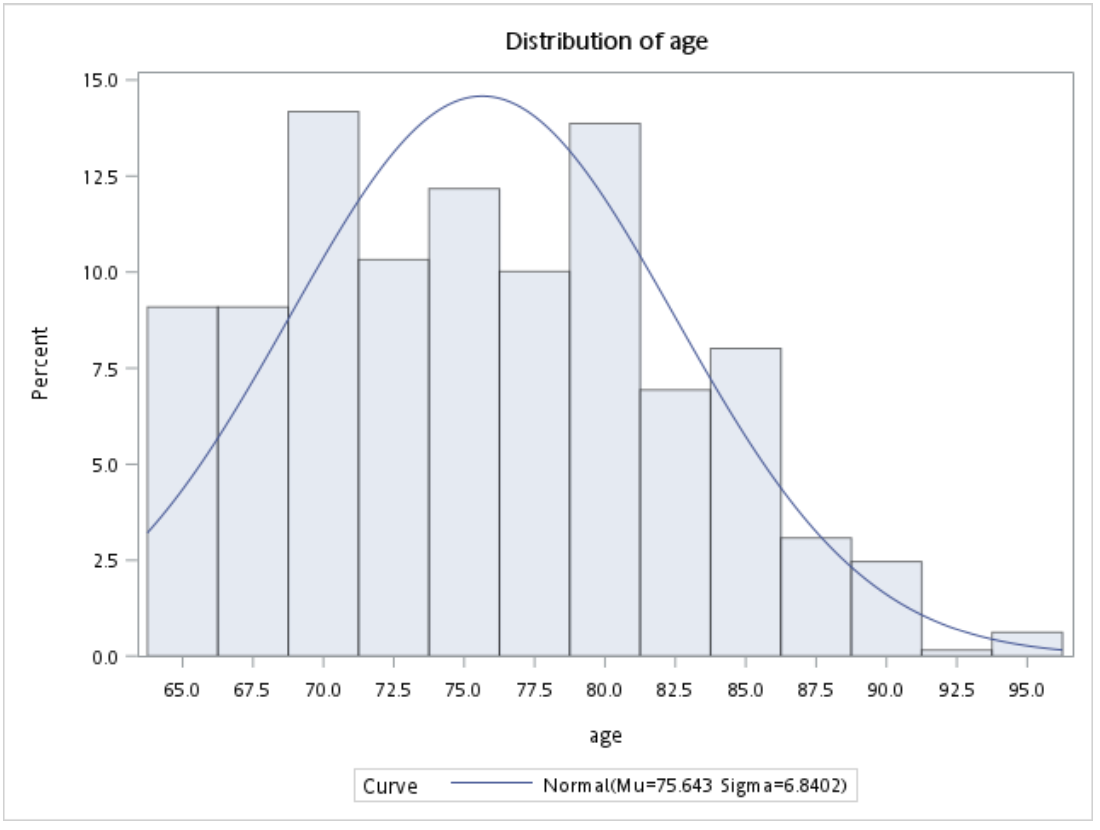
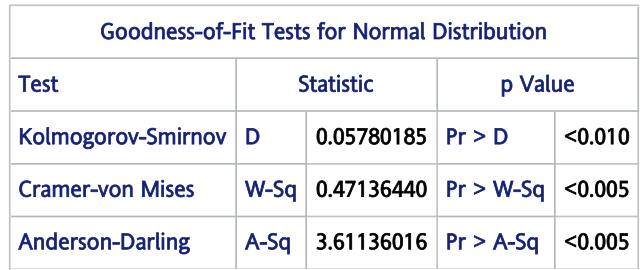
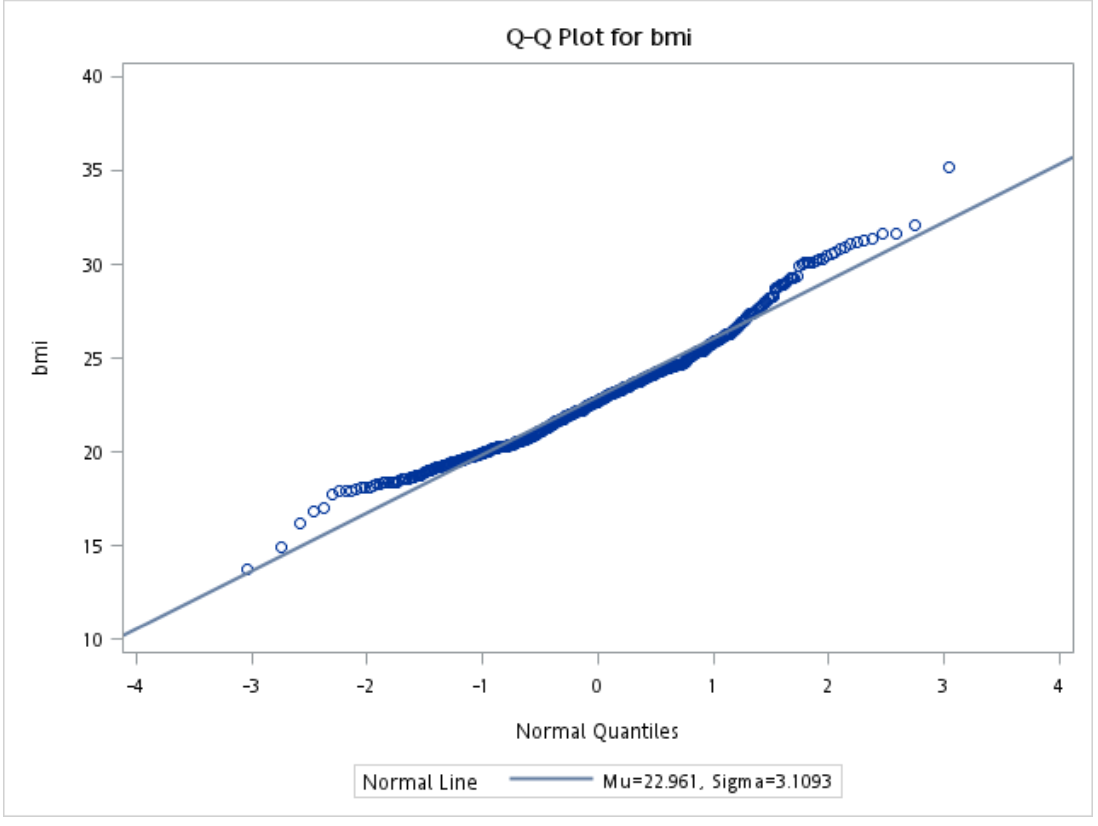
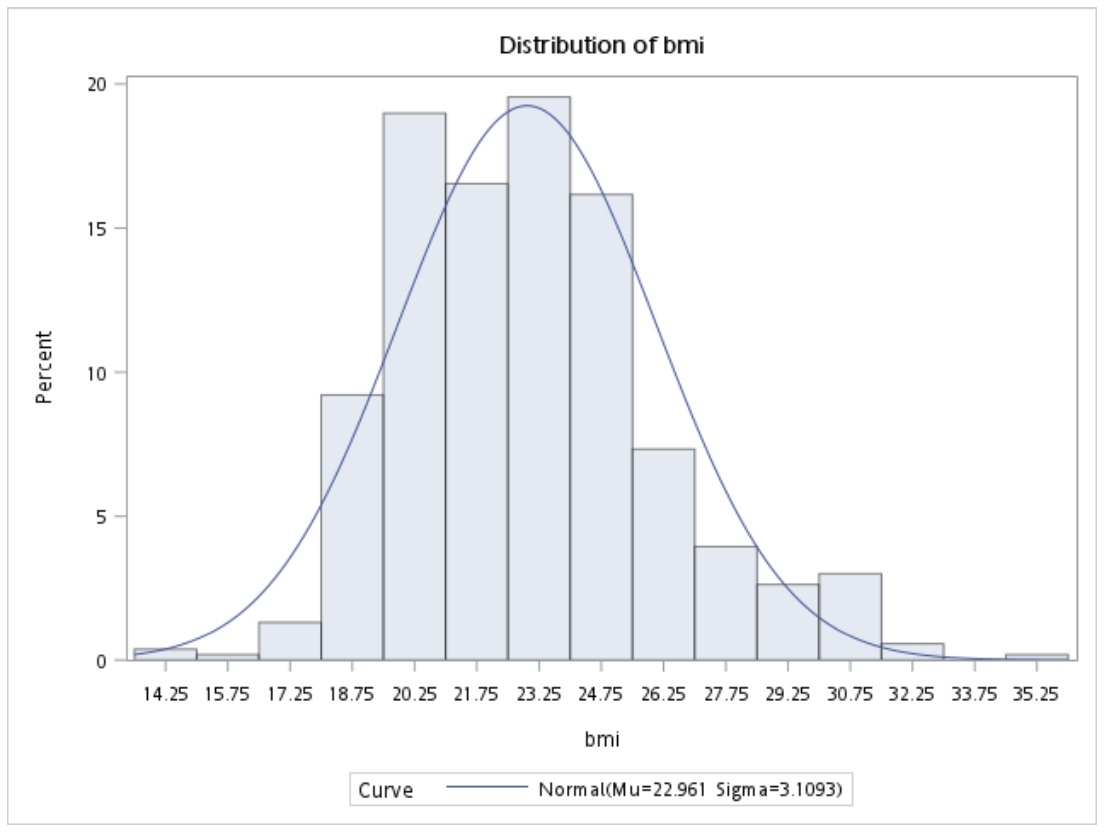
* SAS : homework 9：

學號：b07401048 醫學五 姓名：賴柏瑞

* 1. Univariable analysis of age and BMI
     1. Age：
        1. Result：
           1. 以histogram大致觀察age這個變項的分佈，大致符合常態分佈
           2. 以Q-Q plot大致觀察age這個變項分佈，樣本點大致落於斜線附近，僅在age較小部分稍微偏離。
           3. 以goodness-of-fit test檢定該變項是否符合常態分佈。檢定結果拒絕虛無假說，表示該變項不符合常態分佈。然而根據課程所述，該檢定過於嚴格，容易受到樣本數影響。
        2. Figure：



* + 1. BMI：
       1. Result：
          1. BMI為自訂變項，利用原始資料中的height與weight計算而得。公式為。
          2. 以histogram大致觀察BMI這個變項的分佈，大致符合常態分佈
          3. 以Q-Q plot大致觀察BMI這個變項分佈，樣本點大致落於斜線附近，僅在極端值部分稍微偏離。
          4. 以goodness-of-fit test檢定該變項是否符合常態分佈。檢定結果拒絕虛無假說，表示該變項不符合常態分佈。然而根據課程所述，該檢定過於嚴格，容易受到樣本數影響。
       2. Figure：



* + 1. Code：

*/\*homework 9\*/*

dm "odsresult" clear;

dm "log"clear;

*/\* import data of sasdataset\*/*

libname data "\\Mac\Home\Desktop\";

*/\* q1 \*/*

title "q1";

*/\* univariable analysis (age, bmi)\*/*

**data** ad;

set data.ad\_dataset\_new;

bmi = wt40 / (ht / 100) \*\* 2;

**run**;

**proc** **univariate** data = ad ;

var age bmi;

histogram age bmi / normal;

qqplot age bmi / normal ( mu = est sigma = est);

**run**;

*/\* excluding unsuitable data \*/*

**proc** **print** data = ad;

where bmi > 60;

var ht wt40 bmi age;

**run**;

**data** ad;

set data.ad\_dataset\_new;

bmi = wt40 / (ht / 100) \*\* 2;

if bmi < 60;

**run**;

**proc** **univariate** data = ad ;

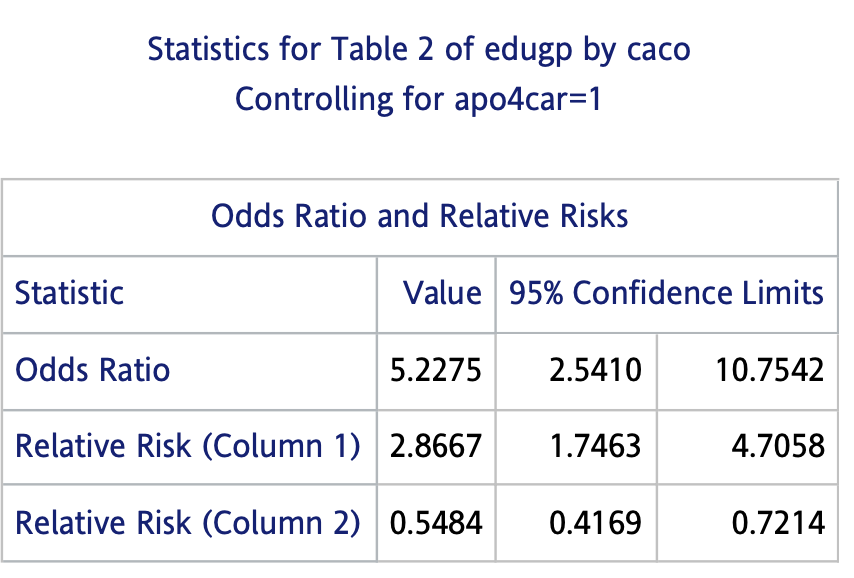
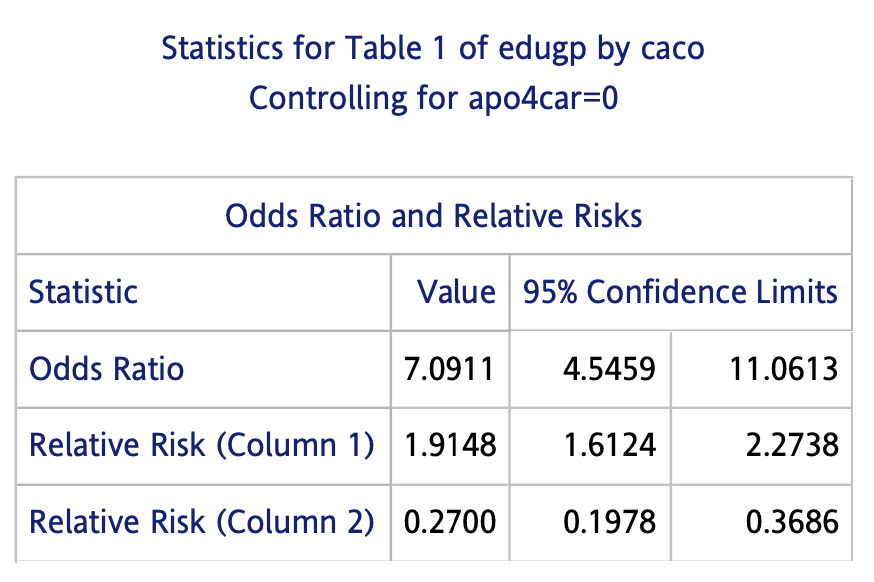
var age bmi;

histogram age bmi / normal;

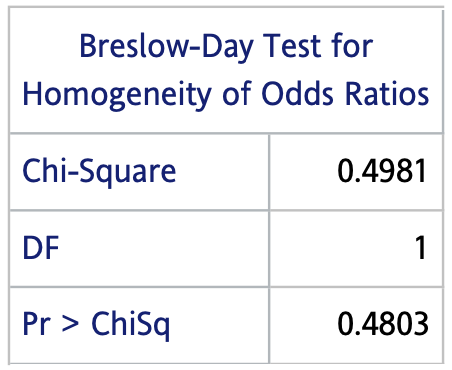
qqplot age bmi / normal ( mu = est sigma = est);

**run**;

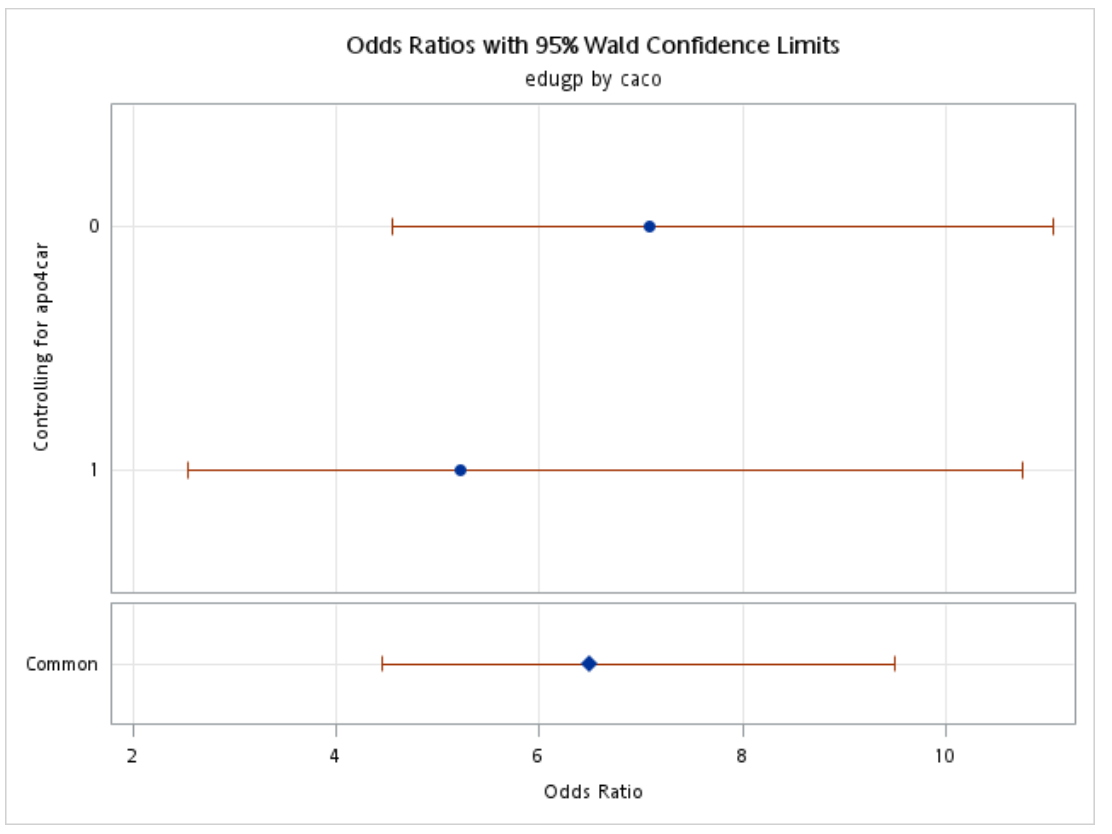
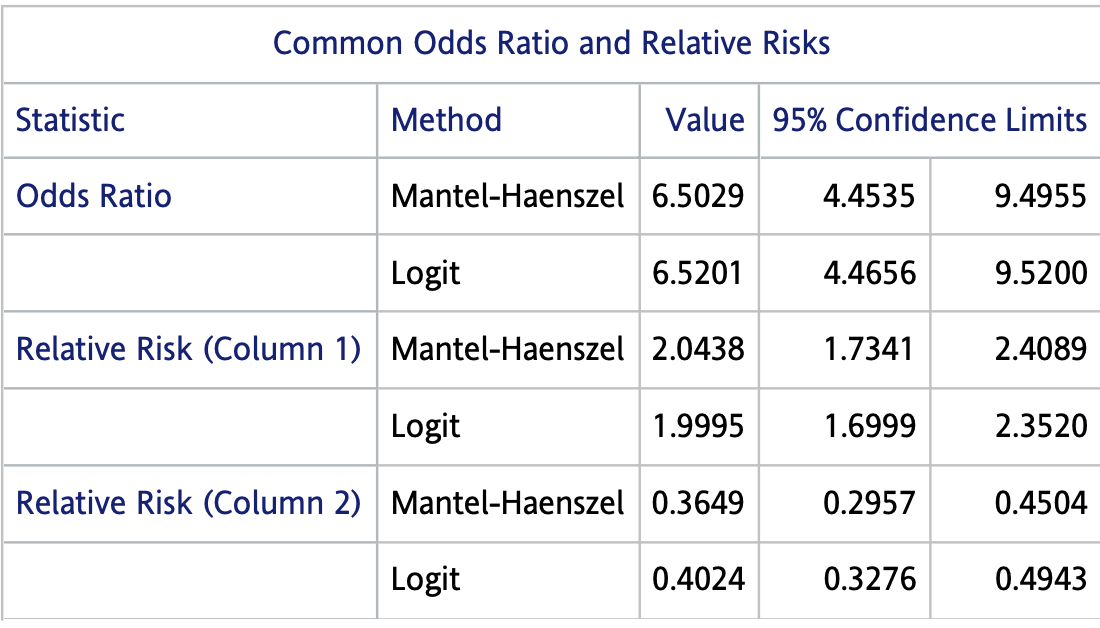
* 1. E ( education status) 🡪 DZ ( AD ), stratified by APOE e4 status
     1. Stratum specific ORs
        1. Result：
           1. Non-APOE e4 carrier stratum specific OR：7.0911 ( 95% CI：4.5459 – 11.0613)
           2. APOE e4 carrier stratum specific OR：5.2275 ( 95% CI : 2.5410 – 10.7542)
        2. Figure：



* + - 1. Descriptions
         1. 將educational status分為小於12年與大於等於12年分為兩組，並定義小於12年為exposure，大於等於12年為non-exposure。
         2. 以是否帶有APOE e4 （potential confounder）將所有樣本分層，觀察兩層中educational status（exposure）與發生AD（disease）之間的關係。
         3. 在有APOE e4的層中的odd ratio為5.2275（95% CI：2.5410 – 10.7542）。在抽樣與暴露相互獨立的前提下，可由OR推論RR。故在有APOE e4層中，education小於12年的組發生AD的風險為education大於12年的5.2275倍，且達統計上顯著。
         4. 在沒有APOE e4的層中的odd ratio為7.0911 ( 95% CI：4.5459 – 11.0613) 。在抽樣與暴露相互獨立的前提下，可由OR推論RR。故在沒有APOE e4層中，education小於12年的組發生AD的風險為education大於12年的7.0911倍，且達統計上顯著。
    1. Breslow-Day homogeneity test
       1. Result：
          1. P-value為0.4803，無法拒絕虛無假說，兩層OR沒有顯著差異。
       2. Figure：



* + - 1. Descriptions
         1. 若要以Mantel-Haenszel計算pooling的OR，需達成兩層OR沒有顯著差異的前提。（”The effect being estimated is constant across the strata”）
         2. 故以Breslow-Day test檢定兩層的同質性。Breslow-Day test的虛無假說為兩層OR相等。
         3. 檢定結果p-value為0.4803，未達統計上顯著，無法拒絕虛無假設，表示兩層OR沒有顯著差異。
    1. Mantel-Haenszel OR
       1. Result：
          1. Common OR：6.5029（95% CI：4.4535 – 9.4955）。
       2. Figure：



* + - 1. Descriptions：
         1. 前題已經以Breslow-Day test檢定兩層具有同質性，故本題以Mantel-Haenszel test計算pooling的OR。
         2. Mantel-Haenszel 檢定結果common odds ratio為6.5029（95% CI：4.4535 – 9.4955）。信賴區間不包含1，達統計上顯著。
         3. 未分層前crude OR為5.8989，分層後unbiased OR 為 6.5029。1 < crude OR < unbiased OR，可見 bias toward null。
         4. 此外，。根據10% rule，APOE e4 為一個confounder。
    1. Code：

*/\* q2 educational status (E) --> AD (DZ), stratified by APOE status \*/*

title "q2";

*/\* grouping educational status , <12 and >=12\*/*

**proc** **univariate** data = ad;

var eduyr;

histogram eduyr / normal;

**run**;

**data** ad2;

set ad;

if eduyr = . then edugp = .;

else if eduyr < 12 then edugp = 1;

else edugp = 0;

**run**;

**proc** **print** data = ad2;

var eduyr edugp;

**run**;

*/\* crude OR \*/*

**proc** **freq** data = ad2;

tables edugp \* caco / or;

**run**;

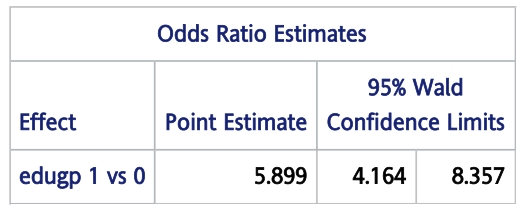
*/\* stratum specific OR \*/*

**proc** **freq** data = ad2;

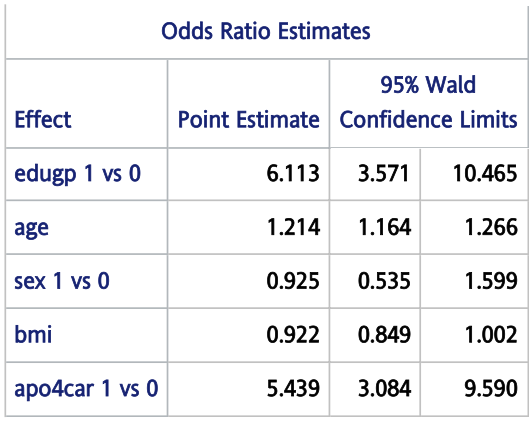
tables apo4car \* edugp \* caco / cmh or ;

**run**;

* 1. E ( education status) 🡪 DZ ( AD ), logistic regression
     1. Unadjusted logistic regression ：
        1. Result：
           1. OR：5.899（95% CI：4.164 – 8.354）
        2. Figure：



* + - 1. Descriptions：
         1. 本題以logistic regression計算結果基本上與上題未分層的crude OR是一樣的。
         2. Education小於12年的人發生AD的機會機會是education大於12年的人的5.899倍。
    1. Adjusted logistic regression ：
       1. Result：
          1. OR：6.113（95% CI：3.571 – 10.465）
       2. Figure：



* + - 1. Descriptions：
         1. 本題為控制age, sex, BMI與APOE e4 status（四個potential confounder）之後進行logistic regression。
         2. Education小於12年的人發生AD的機會機會是education大於12年的人的6.113倍。
         3. 1（H0）< 5.899（unadjusted OR）< 6.113（adjusted OR），可見bias toward H0。
    1. Code：

*/\* logistic regression educational status (E) --> AD(DZ) \*/*

title "q3" ;

*/\* unadjusted \*/*

**proc** **logistic** data = ad2;

class edugp (ref = "0") /param = ref;

model caco (event = "1") = edugp;

**run**;

**proc** **logistic** data = ad2;

class edugp (ref = "0") sex (ref = "0") apo4car (ref = "0") / param = ref;

model caco ( event = "1") = edugp age sex bmi apo4car ;

**run**;

**proc** **logistic** data = ad2;

model caco(event = "1") = eduyr;

**run**;

* 1. In a stratified analysis of data arising in a case-control study, which of the following types of strata contribute information to the Mantel-Haenszel odds ratio:

(a) strata in which all cases and all controls are exposed

(b) strata in which all cases and all controls are unexposed

(c) strata in which all cases are exposed and all controls are unexposed

(d) strata in which there are only controls and no cases