1. For the problem of “Glove Box Door Alignment” experiment to achieve parallelism equal to zero,

1. Calculate the nominal SN ratios (10log10(2/*s*2)) of the parallelisms of the experimental results and all the main and interaction effects on the parallelism SN ratio. Use the third- and fourth-order interaction effects to estimate *seffect* and test which main and two-factor interaction effects are significant (a=0.01);

S/N ratio計算的結果如下：



計算其main effect即兩組、三組、四組參數的interaction effect，計算如下：



使用Excel迴歸功能，可驗證係數結果是一致的。



選擇三次及四次項的interaction effect共五項，算出標準差，可得到：

以及t value

利用及t value進行T test，判定若p value<0.01即表示該effect為顯著的項。T test結果如下：

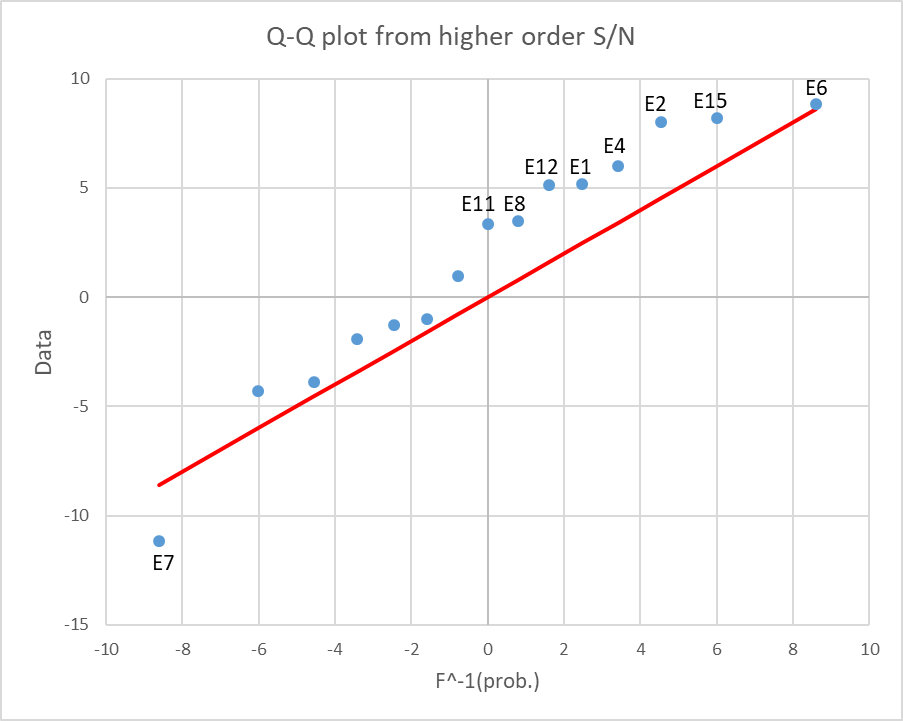


從檢定結果來看，所有的effect皆為不顯著。

1. With the *seffect* and assuming no effects, plot the Q-Q plot of the factor effects and visually determine which factors may have effects on the SN ratio;

將係數排序後，可繪出Q-Q plot如下圖所示。其中對應的常態分佈模型為N(0,4.69023)。





從圖中可以看見，雖然有標記的幾點在T test中由於P value>0.01而被判定為非顯著，但從Q-Q plot中仍可看到該點偏離45度線非常遠。考量到t distribution的參數中，樣本數僅有5，且使用高項次的Effect取標準差會使標準差非常大，加上信心水準有99%，會確實的篩選掉許多顯著但未能通過檢定的項目。因此透過Q-Q plot則可以彌補在這裡T test的盲點。從Q-Q plot中，指定E1, E2, E4, E6, E7, E8, E11, E12, E15。

1. Build a regression model on the parallelism SN ratio using factors with significant effects;

根據上述所選擇的顯著項建立Regression model為

1. Use the model built in (c) to determine the optimum settings to minimize the variability in the parallelism and then use the model on page 61 of SPCO6 to minimize the parallelism to zero.

對上述的model進行最佳化求解，可得到在時，S/N ratio有最大值24.538818，作為最小化實驗分佈的最佳解

對於針對average所建立的model為

所得到的最佳解為當[x1,x2,x3]=[1,1,-1]or[-1,-1,1]時會使y最接近0的解

在兩個階段中的最佳化求解中，是沒有辦法使兩者同時到最小變異及目標值0的狀態，因此需要做取捨。我們將S/N ratio最大的參數組合[1,1,1,1]固定x4, 再選擇的兩個組合中可以使S/N ratio達到最小的參數組合，可以得到

1. Repeat (a) to (c) to build a regression model on the **smaller-the-better SN ratio** (-10log10(S*Xi*2/*n*)) of the parallelism. Determine the optimum settings to maximize the smaller-the-better SN ratio and compare the results to (d)  
     
   S/N ratio計算的結果如下：



計算其main effect即兩組、三組、四組參數的interaction effect，計算如下：



使用Excel迴歸功能，可驗證係數結果是一致的。



選擇三次及四次項的interaction effect共五項，算出標準差，可得到：

以及t value

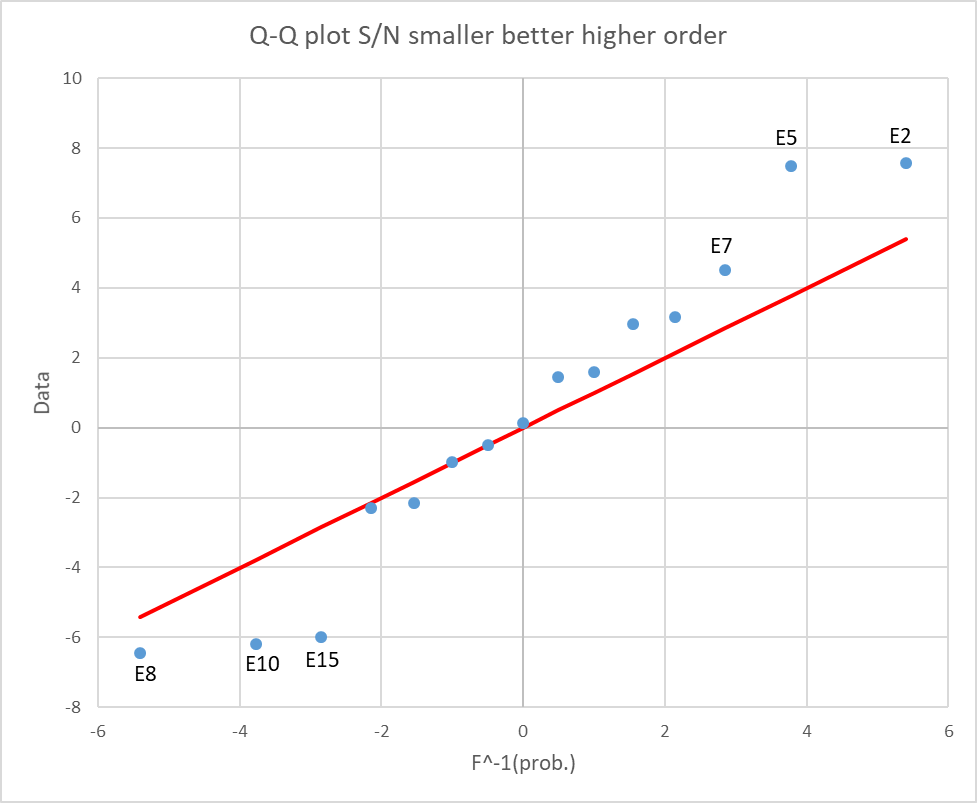
利用及t value進行T test，判定若p value<0.01即表示該effect為顯著的項。T test結果如下：



從檢定結果來看，所有的effect皆為不顯著。

將係數排序後，可繪出Q-Q plot如下圖所示。其中對應的常態分佈模型為N(0,2.9467)。





從圖中可以看見，雖然有標記的幾點在T test中由於P value>0.01而被判定為非顯著，但從Q-Q plot中仍可看到該點偏離45度線非常遠。考量到t distribution的參數中，樣本數僅有5，且使用高項次的Effect取標準差會使標準差非常大，加上信心水準有99%，會確實的篩選掉許多顯著但未能通過檢定的項目。因此透過Q-Q plot則可以彌補在這裡T test的盲點。從Q-Q plot中，指定E2, E5, E7, E8, E10, E15為顯著項。

根據上述所選擇的顯著項建立Regression model為

對其結果進行最佳化求解，可得到在時有最大值。

比較結果與(d)的相異之處，會發現在x3這個參數的高低水準在不同的regression model裡面是不一樣的，可見在這個題目裡面，兩階段的最佳化結果和一次的是不同的。

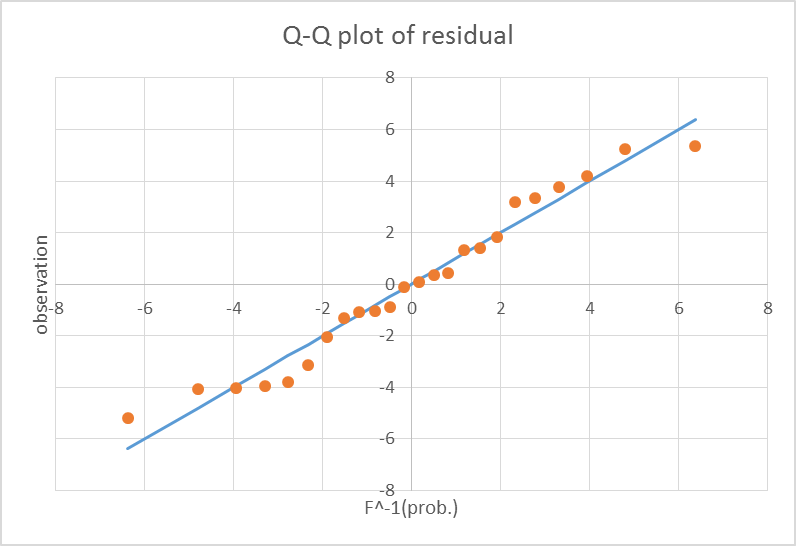
2. For the yield problem in Problem 4 of hw#10,

1. With the model built in (d) of HW#10 Problem 4, perform the residual analysis including Q-Q plot, residual plots vs. run order, , , *xi*.  
   Perform Bartletts’ test with a=0.1 for the variance of the replicates.

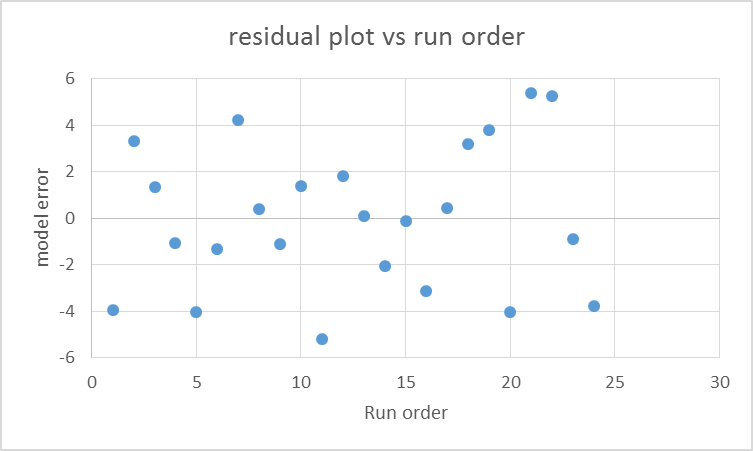
將所建立出來的regression model帶入與每組實驗設定的相同的參數組合，帶入model中求出對系統的預測輸出，將每一組參數組合進行的每次實驗(每個組合都有三組replicates)，分別減去對該組參數組合的預測輸出，可得到其residual。其結果如下：



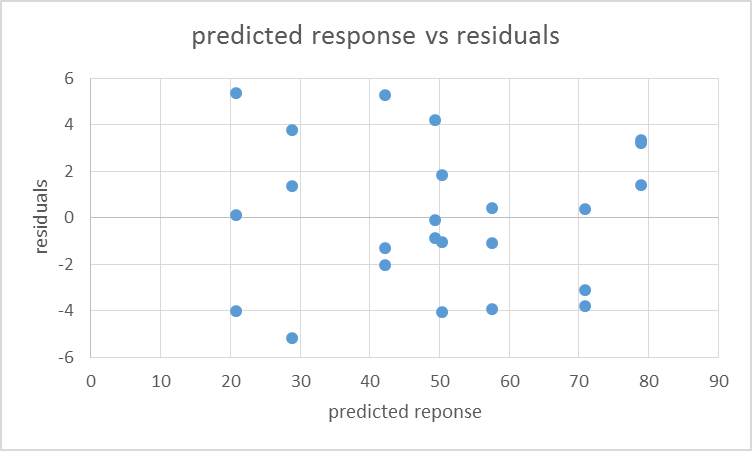
將residual進行分析，可得到其平均為，且樣本標準差為3.131143。將這兩個參數建立常態分佈模型，與data數據比較後，輸出Q-Q plot，結果如下。從結果可看出，residual的資料點與45度線的趨勢接近，可得到residual是符合該常態分佈模型的結論。



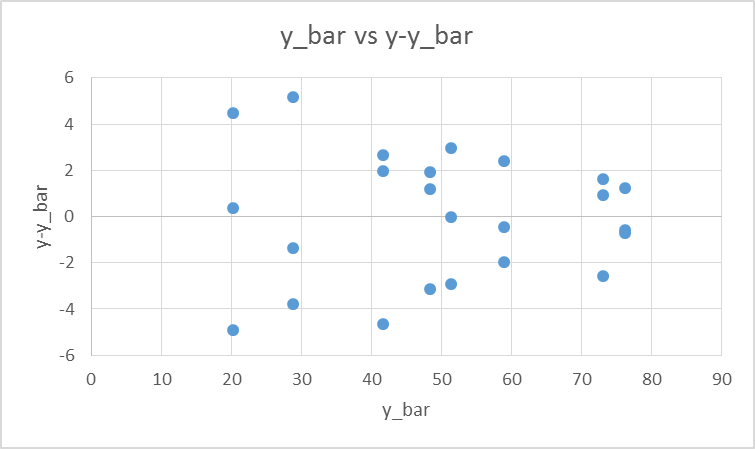
將其run order對residual的關係圖輸出為residual plot vs run order，可得到關係圖如下。從圖中也可看見，residual的出現是隨機的，並不會隨著run order而改變，從結果可知run order對於系統出現residual的影響並不大。



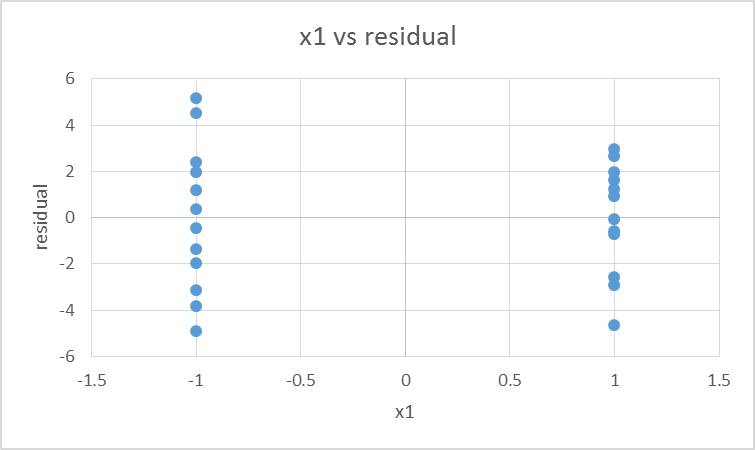
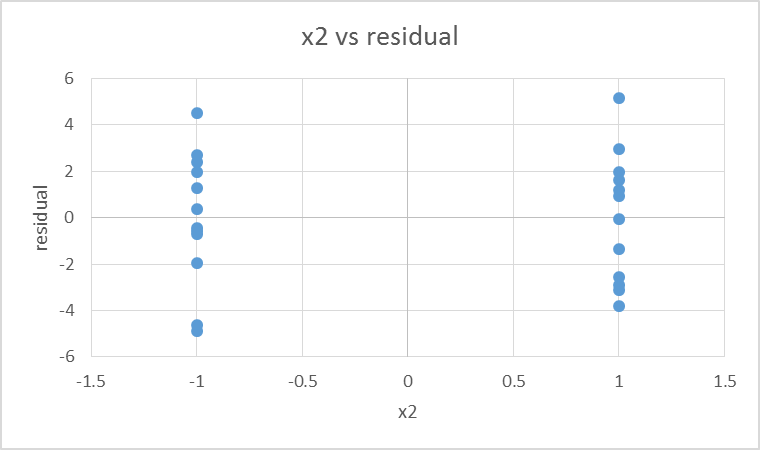
將從模型預測出來的輸出結果與每組實驗所得到的誤差進行關係圖的輸出，可得到關係圖如下。從圖中也可看出模型預測與residual的關係是隨機的，並不會有直接相關，也可以說明預測模型可以把雜訊所造成的偏差去除在模型之外。

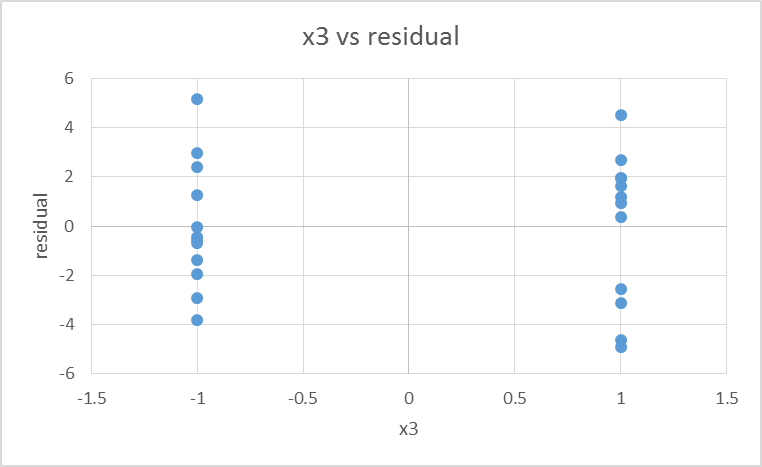


將每一組參數組合實驗的replicates進行平均，作為每一組參數組合的實驗輸出。求出每一組實驗與該組實驗平均值得誤差後，輸出每組實驗的平均值對應到每組實驗的誤差值的關係圖。從圖中可看出，當實驗輸出值越高的時候，誤差分佈的範圍較小的趨勢。



針對、與的對於誤差分佈的影響，將分別將三組參數在高低水準時的residual關係圖plot出來，從結果可以看到，除了的高低水準對於系統的誤差有些許的影響，其他兩者的高低水準表現皆不會對系統有顯著影響。





利用Bartlett’s Test來檢測是否可以將每一組實驗的variance進行等效的判定。建立一假設檢定：

使用的檢定值Test statistic

如果，則推翻無效假設。

從每組實驗的標準差中可以得到M=5.205798，c=1.1875，在90%的信心水準下，可計算，而。

由於，因此沒有充足的理由可以推翻無效假設，可結論為所有的實驗標準差可視為相同的。

1. Calculate the nominal SN ratios (10log10(μ2/*s*2)) of the yield of the experimental results and all the main and interaction effects on the yield SN ratio. Use the smallest two-factor interaction and the three-factor interaction effects to estimate *seffect* and test which main and two-factor interaction effects are significant (a=0.01);

計算每一組實驗的平均值與變異數，並求出nominal S/N ratio



根據S/N ratio計算每一種組合對於S/N ratio的Effect:



進一步以Excel迴歸功能進行驗證，可得到一致的結果：



從表中選出在two-factor interaction中影響最小的和three-factor interaction的effect 計算標準差，所得到的結果為：

計算在DOF=2，信心水準為99%的雙尾檢定t value:

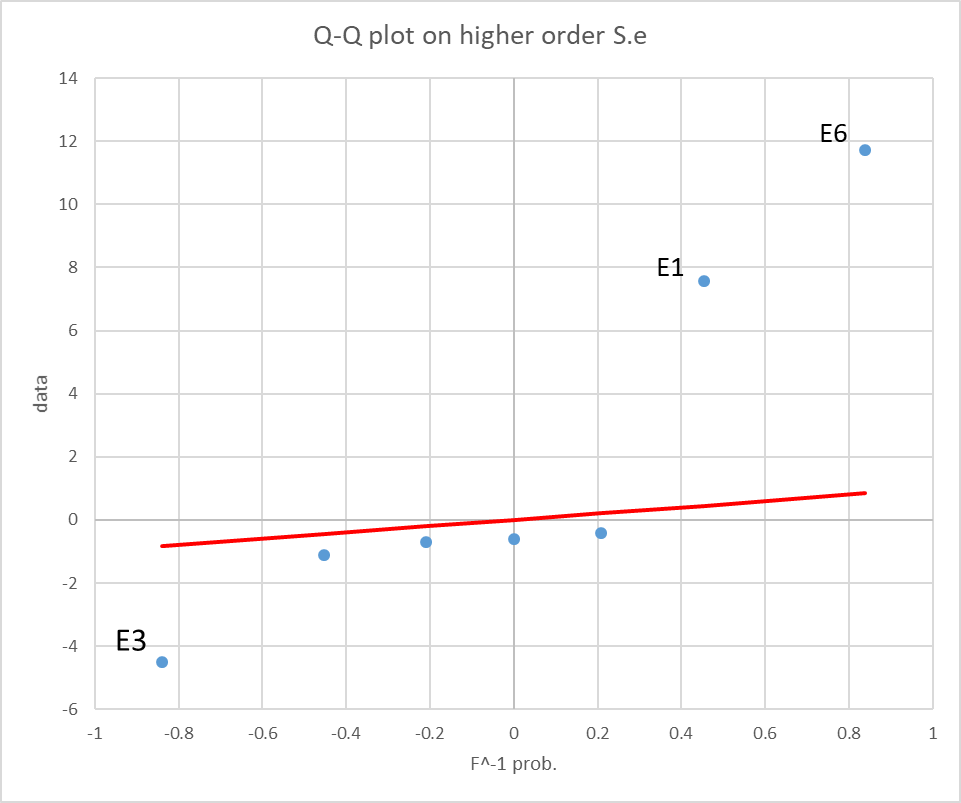
由透過對每個effect進行T檢定，找出相對影響是顯著的main effect及interaction effect，結果如下表。選擇p value小於0.01的項判定為影響顯著的項目，可判定有E1(X1)及E7(X1X2X3)。



1. With the *seffect* and assuming no effects, plot the Q-Q plot of the factor effects and visually determine which factors may have effects on the SN ratio;

將係數排序後，可繪出Q-Q plot如下圖所示。其中對應的常態分佈模型為N(0,0.572252)。





從圖中可以看見，雖然E3在T test中由於P value>0.01而被判定為非顯著，但從Q-Q plot中仍可看到該點偏離45度線非常遠。考量到t distribution的參數中，樣本數僅有2，且使用高項次的Effect取標準差會使標準差非常大，加上信心水準有99%，會確實的篩選掉許多顯著但未能通過檢定的項目。因此透過Q-Q plot則可以彌補在這裡T test的盲點。從Q-Q plot中，除了判定原本就判定為顯著的E1與E6之外，也判定E3(X3)為顯著的影響項目。

1. Build a regression model on the yield SN ratio using factors with significant effects;

將(d)中所得到的三項main effect和interaction effect取出，建立數學模型為：

1. Use the model built in (d) to determine the optimum settings that minimize the variability in the yield and then use the model built in (d) of HW#10 Problem 4 to achieve a 100% yield.

使用規劃求解功能，計算出在(d)的model在這三組參數範圍皆在-1到1之間的最佳解，最佳組合在[]=[1,-1,-1]時可得到。

因此可以得知，在two step optimization中，重現經度以[]=[1,-1,-1]時可以使重現精度最高，而要讓目標值最接近100%，回顧HW#10-4中的數學模型

也會在[]=[1,-1,-1]得到最大的實驗平均，因此可以得知[]=[1,-1,-1]為two step optimiztion的最佳解，且在兩階段中不需要調整參數組合。

1. Repeat (b) to (d) to build a regression model on the **SN ratio** (-10log10(S(100−*Xi*)2/*n*)) of the yield. Determine the optimum settings to maximize the SN ratio and compare the results to (e).

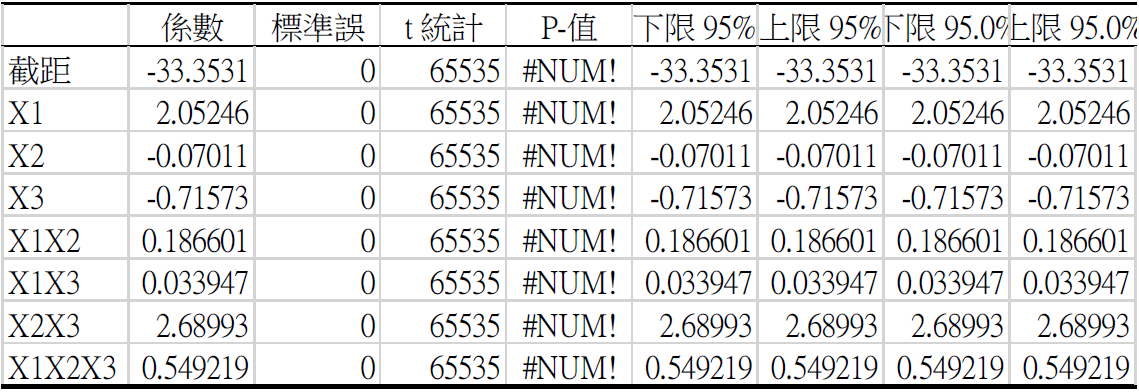
使用「對100距離最小」作為目標S/N ratio，計算其S/N ratio可得到：



計算main effect和interaction effect可得：



使用Excel迴歸工具驗證可得相同結果：



從表中選出在two-factor interaction中影響最小的和three-factor interaction的effect 計算標準差，所得到的結果為：

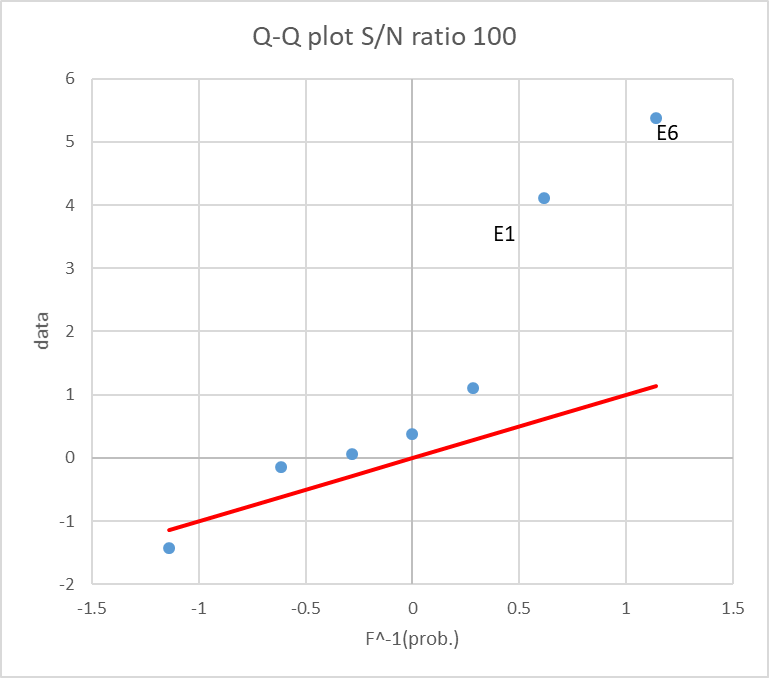
計算在DOF=2，信心水準為99%的雙尾檢定t value:

由透過對每個effect進行T檢定，找出相對影響是顯著的main effect及interaction effect，結果如下表。選擇p value小於0.01的項判定為影響顯著的項目，會發現在這樣的檢定下全部的Effect都被判定為不顯著。



因此透過Q-Q plot，我們可以進一步看到參數的顯著程度。將係數排序後，可繪出Q-Q plot如下圖所示。其中對應的常態分佈模型為N(0,0.77819)。





從Q-Q plot中可看見，E1與E6是相對顯著的。考量到t distribution的參數中，樣本數僅有2，且使用高項次的Effect取標準差會使標準差非常大，加上信心水準有99%，會確實的篩選掉許多顯著但未能通過檢定的項目。因此透過Q-Q plot則可以彌補在這裡T test的盲點。從Q-Q plot中，選擇E1(X1)與E6(X2X3)為顯著項目。

透過兩個顯著的項目，所建立的Regression model為：

使用規劃求解功能，可在[]=[1,-1,-1]與[]=[1,1,1]兩個組合中皆取得最佳S/N ratio 。

與(e)做比較，可發現在[]=[1,-1,-1]時兩者皆能擁有最佳解，唯一不同是由於(f)中的model並未對項進行單調性的限制，因此在進行最佳化演算時會出現多一組解。

使用與100的距離的S/N ratio的概念即是結合了「平均程度」與「對目標的距離」同時最小化的概念，相比於two step optimization所帶出的「先最小化Variance再逼近目標值」來的更直觀，從這次的題目中也可以看到，兩者在結果上是相同的。使用改良過的S/N ratio更好的地方是，若two step optimization中的兩階段所求出來的最佳解參數組合是不一樣的，可能會需要在平均度與對目標距離這兩項目標函數中進行權重的改變，屆時又會有其他的假設存在，而改良過的S/N ratio是針對單一目標，在最佳化的計算中較單純。