**HW 4: Diagnose the validation of Cox PH model for the gastric data**

Gastric cancer data are shown in “gastric cancer data.csv”, where Treatment= 0 if chemotherapy only; 1 if chemotherapy with radiotherapy.

Please answer the following questions and summarize and explain your results. Computer outputs without any interpretations are not acceptable.

1. Please draw Kaplan-Meier curves stratified by treatments and interpret the result.

一張含有 文字, 圖表, 行, 地圖 的圖片

AI 產生的內容可能不正確。

Log-rank test:





Treatment=0（化療）與Treatment=1（化療＋放療）的生存曲線，在1000天左右的位置生存曲線交叉，整體上差異不大。根據 Log-rank 檢定，p值= 0.64 (> 0.05)，兩組之間的生存差異不顯著。因此，我們沒有足夠的證據認為放射治療能顯著改變病人的生存時間。

1. Please fit the data with Cox PH model and interpret the result.







一張含有 文字, 螢幕擷取畫面, 字型 的圖片

AI 產生的內容可能不正確。

根據 Cox 比例風險模型的結果：治療方式（化放療 vs. 化療）對病人存活時間的影響不顯著。化放療組的估計風險比 (HR = 1.11) 稍高於化療組，但差異並不顯著（p = 0.638）。表示根據資料，加上放射治療並沒有明顯降低或提高死亡風險。因此，無法拒絕虛無假設，治療方式沒有顯著影響生存風險。

1. Following Question 2, draw the fitted survival curve on the same figure of Question 1 and compare with Kaplan-Meier curves. Use the result to conduct model checking.

一張含有 文字, 圖表, 行, 螢幕擷取畫面 的圖片

AI 產生的內容可能不正確。

Cox 擬合曲線與 KM 曲線有明顯差距，有可能是因為Cox hazard ratio隨時間變化，比例風險假設可能不成立，所以Cox曲線擬合得不好。

1. Use the log-log plots to diagnose Cox model.

一張含有 文字, 圖表, 行, 螢幕擷取畫面 的圖片

AI 產生的內容可能不正確。

從 log-log plot 觀察到兩組曲線不太平行且在後期有交叉，加放療組曲線下降至低於單純化療組，顯示比例風險假設不成立。這與 Cox 擬合KM 曲線距離較遠的現象一致，表示 Cox 模型在此資料中對生存函數的估計不好， Cox 模型用固定危險比描述 Treatment 效果，會錯估後期加放療組的事件風險（生存機會）。

1. Use Schoenfeld residuals for testing time-varying effect, and draw time-varying effect.

一張含有 行, 圖表, 文字, 平行 的圖片

AI 產生的內容可能不正確。

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

AI 產生的內容可能不正確。

Treatment:





GLOBAL:





Schoenfeld 殘差檢定顯示 Treatment 的比例風險假設不成立（p < 0.001），表示其危險比隨時間變化；同時也違反整體模型 PH的假設。

程式碼網址:

<https://github.com/Lai-jun-yan/Survival_data_analysis/tree/master/%E4%BD%9C%E6%A5%AD3>