**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, 地圖, 圖表, 螢幕擷取畫面 的圖片

AI 產生的內容可能不正確。

Treatment=0（化療）與Treatment=1（化療＋放療）的生存曲線在整體上幾乎重疊。p值= 0.64 (> 0.05)，根據 Log-rank 檢定，兩組之間的生存差異不顯著。因此，從統計上看，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）。 從臨床角度來看，這表示在本資料中，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「兩組生存率相同」的虛無假設

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，且 Cox hazard ratio 似乎隨時間變化，顯示比例風險假設可能不成立，Cox 擬合曲線擬合得不好。

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

一張含有 文字, 圖表, 行, 字型 的圖片

AI 產生的內容可能不正確。

從 log-log plot 觀察到兩組曲線在後期有交叉，顯示比例風險假設不成立。這與 Cox 擬合曲線與 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時間的擬合欠佳，特別是後期的事件風險可能被低估或高估。

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

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。

一張含有 文字, 行, 圖表, 斜率、斜坡 的圖片

AI 產生的內容可能不正確。

使用Schoenfeld殘差檢驗Cox模型中Treatment的比例風險假設，結果顯示 p = 0.00028，Treatment效應隨時間變化；GLOBAL p = 0.00028 < 0.05 ，整體模型比例風險假設也違反恆定。β(t) 曲線顯示，Treatment 效應隨時間下降。曲線左上方表示早期 hazard ratio 較高，右下方表示後期 hazard ratio 較低，橫跨紅色水平線，比例風險假設不成立。

程式碼網址:

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