**Bài tập thực hành số 6:**

**Phân tích sống**

**Thời gian phục hồi vận động sau phẫu thuật thay khớp háng ở người cao tuổi.**

**Giới thiệu**

Nghiên cứu này được tiến hành để theo dõi thời gian cần thiết để bệnh nhân từ 60 tuổi trở lên đạt được tình trạng vận động hoàn toàn bình thường sau phẫu thuật thay khớp háng, và các yếu tố ảnh hưởng. Nghiên cứu được tiến hành với 1600 bệnh nhân phẫu thuật thay khớp háng, thời gian theo dõi là 12 tháng.

Các biến số trong bộ số liệu (**data6.csv**) như sau

**Biến phụ thuộc:**

* thoi\_gian\_phuc\_hoi: Thời gian (tính bằng tháng) từ khi phẫu thuật đến khi đạt được tình trạng vận động hoàn toàn bình thường hoặc bị kiểm duyệt.
* trang\_thai\_su\_kien : Biến nhị phân chỉ trạng thái vào cuối thời gian theo dõi/kiểm duyệt:
  + 1: Đã đạt được tình trạng vận động hoàn toàn bình thường (sự kiện xảy ra).
  + 0: Đã bị kiểm duyệt (censor) do: tử vong trước khi phục hồi, mất dấu trước 12 tháng, hoặc đạt đến 12 tháng theo dõi mà chưa phục hồi.

**Biến độc lập:**

1. tuoi : Tuổi của bệnh nhân khi phẫu thuật (liên tục, từ 60 trở lên)
2. gioi : Giới tính (phân loại: "Nam", "Nữ")
3. tinh\_trang\_kinh\_te : Tình trạng kinh tế (phân loại: "Nghèo", "Trung Bình", "Cao")
4. tinh\_trang\_nha\_o : Tình trạng nhà ở (phân loại: "Sống một mình", "Sống với người thân", "Khác")
5. mac\_benh\_khac : Có mắc bệnh khác không (phân loại: "Có", "Không")
6. thoi\_gian\_nam\_vien : Thời gian nằm viện sau phẫu thuật (liên tục, tính bằng ngày)
7. ho\_tro\_phuc\_hoi\_CN : Có nhận được hỗ trợ phục hồi chức năng sau phẫu thuật không (phân loại: "Có", "Không")
8. chi\_so\_dinh\_duong : Chỉ số đánh giá tình trạng dinh dưỡng (liên tục, thang điểm 1-10)
9. so\_lan\_tai\_kham : Số lần tái khám trong 3 tháng đầu sau phẫu thuật (biến đếm)
10. bien\_chung\_som : Có biến chứng sớm trong tháng đầu sau phẫu thuật không (phân loại: "Có", "Không")

Nhiệm vụ

1. Vẽ đường cong sống Kaplan-Meier chung và theo nhóm (ví dụ: gioi, ho\_tro\_phuc\_hoi\_CN).
   * Dựa vào đường cong chung, ước tính tỷ lệ bệnh nhân đạt được vận động bình thường sau 6 tháng và 12 tháng.
   * Có sự khác biệt rõ rệt về thời gian phục hồi giữa Nam và Nữ không? p-value từ Log-rank test cho bạn biết điều gì?
   * Những bệnh nhân có/không có hỗ trợ phục hồi chức năng có thời gian phục hồi khác nhau không? Giải thích ý nghĩa p-value.
2. Xây dựng mô hình Cox đa biến để xác định các yếu tố dự đoán thời gian phục hồi
   * Những biến độc lập nào có ý nghĩa thống kê (p-value < 0.05) trong việc dự đoán thời gian phục hồi?
   * Giải thích ý nghĩa của Hazard Ratio (HR) cho biến tuoi. (Gợi ý: HR < 1 nghĩa là giảm nguy cơ sự kiện, HR > 1 nghĩa là tăng nguy cơ sự kiện).
   * Giải thích ý nghĩa của Hazard Ratio cho ho\_tro\_phuc\_hoi\_CNCo. Điều này có nghĩa là gì đối với những bệnh nhân có hỗ trợ phục hồi chức năng so với những người không có?
   * Biến mac\_benh\_khac có HR là bao nhiêu? Diễn giải nó.
3. Kiểm tra giả định bằng cách sử dụng biểu đồ Schoenfeld residuals và kiểm định thống kê.

* Có biến nào vi phạm giả định không? Nếu có, bạn sẽ làm gì để xử lý (ví dụ: thêm tương tác với thời gian, sử dụng mô hình Cox mở rộng, chia thời gian theo dõi thành các khoảng, v.v.)?