

Lista 12

1. Wczytaj dane HF, w których znajdują się następujące informacje: wiek, cukrzyca (1 – tak, 0 – nie), frakcja wyrzutowa, wysokie ciśnienie krwi, kreatynina w surowicy, sód w surowicy, płeć (1 – M, 0 – K), palenie (1 – tak, 0 – nie), czas (miesiące), zdarzenie.
2. Stwórz tabele licznosci dla kolumn, anemia, cukrzyca, płeć i palenie.
3. Do czego służy funkcja `get_dummies`, znajdująca się w pandasie. Użyj jej, aby przetransformować zmienne katagoryczne, usuń z danych oryginalne kolumny.
4. Stwórz model Coxa, który najlepiej opisze przeżywalność w analizowanym zbiorze danych. Zinterpretuj otrzymany model.
5. Sprawdź, czy zostało spełnione założenie o proporcjonalnym hazardzie.
6. Co zwraca funkcja `baseline_hazard_`?
7. Stwórz i zinterpretuj wykres, na który będzie logarytm współczynnika ryzyka dla każdej ze zmiennych.
8. Używając funkcji `plot_partial_effects_on_outcome`, sprawdź, jak zmienia się funkcja przeżycia w zależności od tego, czy pacjent ma 40, 60, czy 95 lat.
9. Stwórz nowe dane, używając poniższych informacji:
 - a. Pacjent 1: wiek – 65, frakcja wyrzutowa – 35, kreatynina w surowicy – 1.1, wysokie ciśnienie krwi – 1
 - b. Pacjent 2: wiek – 72, frakcja wyrzutowa – 25, kreatynina w surowicy – 2.0, wysokie ciśnienie krwi – 0
 - c. Pacjent 3: wiek – 80, frakcja wyrzutowa – 20, kreatynina w surowicy – 1.8, wysokie ciśnienie krwi – 1
10. Użyj przygotowanych w poprzednim zadaniu danych do wyznaczenia mediany czasu przeżycia i funkcji przeżycia dla trzech pacjentów.