

Rozszerzone Modele Liniowe

Regresja Poissona

Zainstaluj pakiet *datasets*.

1. Zbiór danych *warpbreaks* zawiera dane ilustrujące zależność liczby pęknięć w osnowie przędzy w funkcji typu wełny (typy A i B) i siły naprężenia (Low (L), Medium(M), High (H)).
 - a) Porównaj boxploty liczby pęknięć dla obu typów wełny.
 - b) Utwórz zmienną $Y = \log(breaks)$ i porównaj boxploty zmiennej Y dla obu typów wełny.
 - c) Przetestuj hipotezę, że typ wełny ma wpływ na wartość oczekiwaną zmiennej Y za pomocą testu Studenta.
 - d) Porównaj boxploty liczby pęknięć dla różnych wielkości siły naprężenia.
 - e) Porównaj boxploty zmiennej Y dla różnych wielkości siły naprężenia.
 - f) Zbadaj zależność zmiennej Y od zmiennych *wool* i *tension* za pomocą regresji wielokrotnej (procedura *lm*). Podaj wyniki testów istotności.
 - g) W oparciu o powyższą analizę oszacuj oczekiwaną liczbę pęknięć dla przędzy z wełny typu A poddanej średniemu naprężeniu.
 - h) Przeanalizuj zależność liczby pęknięć od typu wełny za pomocą regresji Poissona. Podaj wyniki testów istotności i porównaj je z wynikami uzyskanymi w punkcie f). Które wyniki wydają się bardziej wiarygodne ?
 - i) W oparciu o regresję Poissona oszacuj oczekiwaną liczbę pęknięć dla przędzy z wełny typu A poddanej średniemu naprężeniu. Porównaj z wynikiem osiągniętym w punkcie g).
 - j) Przeanalizuj zależność liczby pęknięć od typu wełny za pomocą regresji dla rozkładu ujemnego dwumianowego. Podaj wyniki testów istotności i krytycznie porównaj je z wynikami uzyskanymi w punktach f) i h).
 - i) W oparciu o regresję dla rozkładu ujemnego dwumianowego oszacuj oczekiwaną liczbę pęknięć dla przędzy z wełny typu A poddanej średniemu naprężeniu. Porównaj z wynikami osiągniętymi w punktach g) i i).
 - j) Na jednym wykresie narysuj zależność średniej wartości zmiennej Y od rodzaju naprężenia, osobno dla dwóch rodzajów wełny. W oparciu o ten wykres sformułuj swoje przewidywania dotyczące interakcji między typem wełny a rodzajem naprężenia.
 - k) Przeanalizuj dane za pomocą wszystkich trzech metod (*lm*, *glm*, *glm.nb*) zakładając model z interakcjami. Podaj i porównaj wyniki testów istotności. Wykorzystaj nowe modele do oszacowania oczekiwanej liczby pęknięć dla przędzy z wełny typu A poddanej średniemu naprężeniu. Porównaj z wcześniej uzyskanymi wynikami.
 - k) Porównaj *residual deviance* dla modelu regresji Poissona i regresji dla rozkładu ujemnego dwumianowego. Który model wydaje się być bardziej odpowiedni dla tego zbioru danych ?

Malgorzata Bogdan