

## Statystyka II Mat. Egzamin komputerowy 4.09.12.

1. Niech dany będzie model liniowy  $y = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ . Wykorzystując rozkład SVD macierzy  $\mathbf{X}$  (funkcja `SVD`), bez użycia funkcji `summary`, mnożenia macierzowego i odwracania macierzy, napisz ogólną funkcję obliczającą wartości statystyki t-studenta dla hipotez  $\beta[i]=0$ . Argumentami funkcji mają być  $\mathbf{y}$  oraz  $\mathbf{X}$ .
2. Policz streszczenie modelu poissonowskiej regresji logliniowej, w którym  $\log(\mathbf{E}(\mathbf{y})) = \mathbf{X}\boldsymbol{\beta}$  na danych `Pima=rbind(Pima.te,Pima.tr)` z biblioteki `MASS`, gdzie  $\mathbf{y}=\text{npreg}+1$ . Następnie policz przybliżoną p-wartość testu ilorazu wiarygodności modelu mniejszego `m1` do modelu pełnego. Model `m1` jest oparty na stałej (intercept), `glu`, `age` i `type` oraz spełnia takie ograniczenie, że współczynnik przy `type` jest 3 razy mniejszy od stałej.
3. Niech czas życia żarówki  $C$  ma rozkład gamma z parametrami: skala=1, kształt=3. Wylosuj 1000 niezależnych obserwacji z tego rozkładu i oszacuj p-stwo, że  $C$  jest nie większy niż  $d=1.5*z_{0.75}$ , gdzie  $z_{0.75}$  jest kwantylem próbkowym rzędu 0.75 tego rozkładu (wskazówka: użyj funkcji `quantile`). Policz  $d$  numerycznie (bez losowania).
4. Wczytaj z katalogu <http://www.mimuw.edu.pl/~pokar/StatystykaII/DANE/DaneEgzaminacyjne/archiwum/MacierzePot1.zip> 12 potencjałów kontaktowych (PK) zawierających energię oddziaływania między aminokwasami w białkach. (PK są to symetryczne macierze 20 x 20). Napisz funkcję `upperVec`, która przekształca macierz na wektor złożony z elementów należących do jej górnej części. Policz macierz korelacji (CC) między takimi wektorami utworzonymi z macierzy PK. Czy w zbiorze wektorów  $x_1, \dots, x_{12}$  są obserwacje odstające?
5. Przekształć CC na macierz odległości euklidesowych i wykonaj klasteryzację `average linkage`. Dendrogram zapisz do pliku `.ps`. (PK powinny być na rysunku oznaczone numerami plików, z których zostały wczytane). Ile jest klastrów?