

# TESTOWANIE HIPOTEZ.

## ZAJĘCIA NR 12.

1. (W, 1/160) Które przyczyny z poniższych mogą spowodować że statystyki  $t$  będą niewłaściwe? (tzn. nie będą miały rozkładu  $t$  przy założeniu, że hipoteza zerowa jest prawdziwa)
  - (i) Heteroskedastyczność.
  - (ii) Korelacja z próby równa 0.95 między dwoma zmiennymi objaśniającymi, które są uwzględnione w modelu.
  - (iii) Pominięcie ważnej zmiennej objaśniającej w modelu.

2. (W, 4/160) Czy na wysokość czynszów wpływa udział studentów w populacji miasta uniwersyteckiego? Niech  $rent$  oznacza średniemu, miesięcznemu czynszowi w mieście uniwersyteckim w USA. Niech  $pop$  oznacza całkowitą populację miasta,  $avginc$  oznacza średni dochód w danym mieście oraz  $pctstu$  oznacza wielkość populacji studenckiej w mieście jako procent całkowitej populacji. Dzięki poniższemu modelowi możemy przetestować istnienie relacji:

$$\log(rent) = \beta_0 + \beta_1 \log(pop) + \beta_2 \log(avginc) + \beta_3 pctstu + \varepsilon$$

- (i) Zapisz hipotezę zerową, która mówi że rozmiar populacji studenckiej w stosunku do całkowitej populacji nie ma wpływu, przy innych czynnikach niezmiennych, na miesięczny czynsz. Zapisz hipotezę alternatywną, która mówi że taka relacja istnieje.
- (ii) Jakich znaków spodziewasz się przy  $\beta_1$  oraz  $\beta_2$ ?
- (iii) Równanie z (i) oszacowano przy użyciu danych z 64 miast uniwersyteckich (w nawiasach błędy standardowe):

$$\widehat{\log(rent)} = .043 + .066 \log(pop) + .507 \log(avginc) + .0056 pctstu$$

(.844)            (.039)            (.081)            (.0017)

$n = 64, R^2 = .458$ . Co jest nie tak w stwierdzeniu: 10% wzrost w populacji jest związany z około 6.6% wzrostem w czynszu?

- (iv) Przetestuj hipotezę z (i) na poziomie istotności 1%.
3. (W, 5/160) Oszacowano model który wiąże przypadki opuszczania zajęć na średnią ocen ( $colGPA$  - średnia ocen,  $skipped$  - średnia liczba opuszczonych wykładów w tygodniu,  $hsGPA$  - średnia ocen z liceum, w nawiasach błędy standardowe):

$$\widehat{colGPA} = 1.39 + .412 hsGPA + .015 ACT - .0837 skipped$$

(.33)            (.094)            (.011)            (.026)

$n = 141, R^2 = 0.234$ .

- (i) Używając przybliżenia standardowym rozkładem normalnym, znajdź 95% przedziały dla  $\beta_{hsGPA}$ .
  - (ii) Czy możesz odrzucić hipotezę zerową:  $b_{hsGPA} = .4$  przeciwko obustronnej alternatywie na poziomie istotności 5%?
  - (iii) Czy możesz odrzucić hipotezę zerową:  $b_{hsGPA} = 1$  przeciwko obustronnej alternatywie na poziomie istotności 5%?
4. (W, 2/258) Poniższe równania zostały oszacowane używając danych BWGHT:

$$\widehat{\log(bwght)} = \begin{array}{cccccc} 4.66 & -.0044cigs & +.0093\log(faminc) & +.016parity & +.027male & +.055white \\ (.22) & (.0009) & (.0059) & (.006) & (.010) & (.013) \end{array}$$

$$n = 1388, R^2 = .0472$$

oraz

$$\widehat{\log(bwght)} = \begin{array}{cccccccc} 4.65 & -.0052cigs & +.0110\log(faminc) & +.017parity & +.034male & +.045white & -.0030motheduc & +.0032fatheduc \\ (.38) & (.0010) & (.0085) & (.006) & (.011) & (.015) & (.0030) & (.0026) \end{array}$$

$$n = 1191, R^2 = .0493.$$

- (i) Skomentuj istotność statystyczną zmiennej *motheduc*.
- (ii) Czy korzystając z podanych informacji, jest możliwe przetestowanie łącznej istotności zmiennych *motheduc* oraz *fatheduc*? Wyjaśnij co zrobiłbyś/abyś aby obliczyć statystykę  $F$ ?