

Ćwiczenia ze statystyki 21 maj 2008.

Poniższy program wyznacza moc na podstawie 1000 obserwacji dla testu t-Studenta, poziomu istotności 0.05 w sytuacji, gdy różnica w średnich pomiędzy rozkładami wynosi 0.5 a liczebność próby wynosi 20.

```
decyzje = NULL
for (i in 1:1000) {
  x = rnorm(20)
  y = rnorm(20)+0.5
  decyzje[i] = t.test(x,y)$p.value < 0.05
}
mean(decyzje)
```

Modyfikując ten program sprawdź

1. Czy dla opisanego przypadku większą moc ma test t-Studenta (t.test) czy Wilcoxon (wilcoxon.test)?
2. Co się stanie z mocą jeżeli próby będą liczebniejsze, np. 50 obserwacji? A co dla 500?
3. Co się stanie z mocą jeżeli próby będą pochodziły z rozkładów o różnej wariancji?
4. Co się stanie z mocą jeżeli próby będą pochodziły z rozkładu log-normalnego?
Rozkład log-normalny, to taki rozkład, że logarytm zmiennej o tym rozkładzie ma rozkład normalny.
5. Porównaj moce testu Kołomogorova-Smirnova i Shapiro-Wilka gdy obserwacje pochodzą nie z rozkładu normalnego ale z rozkładu wykładniczego, jednostajnego lub χ^2 .
6. Porównaj moce testu proporcji (prop.test) i testu t-Studenta (t.test) gdy obserwacje pochodzą z rozkładu dwumianowego.

Zadania „na kartke”

1. Chcemy sprawdzić, czy lek X poprawia stan pacjenta, w tym celu przebadano 100 pacjentów i otrzymano wyniki jak w tabeli poniżej. Sprawdź, czy lek X działa.
2. Chcemy sprawdzić, czy w małżeństwach częściej chorują kobiety czy mężczyźni. Przebadano 50 małżeństw (nie w Polsce) i otrzymano wyniki jak w tabeli. Sprawdź, czy któraś płć jest narażona na częstsze zachorowania.
3. Dzieci w przedszkolu dosyć często chorują. Chcemy sprawdzić, czy częstość zachorowań w 3 wybranych przedszkolach jest podobna, czy też w którymś jest wyższa.

stan pacjenta / leczenie	lek X	placebo
bez zmian	20	30
poprawa	30	20

choruje częściej / choruje rzadziej	K	M
K	2	21
M	26	1

zdrowie / przedszkole	A	B	C
było chore	21	28	34
było zdrowe	32	27	13