

# Kognitywistyczne Seminarium Metodologiczne

Czwartek, 17 kwietnia, g. 17:00, Instytut Informatyki UJ, Łojasiewicza 6, s. 1093

referuje

**mgr Michał Świętek**

(Instytut Matematyki UJ)

## Wzory statystyczne kontra *resampling* i moc obliczeniowa

Z praktycznego punktu widzenia klasyczna statystyka posiada bardzo poważną wadę - jej wzory i twierdzenia mają założenia. Problem z założeniami jest dwojaki. Po pierwsze, ze względu na wysoką komplikację wzorów opisujących rozkłady prawdopodobieństwa, tylko niewiele z interesujących miar statystycznych posiada gotowe formuły na ich istotność statystyczną. Po drugie, rzeczywistość nierzadko nie współgra z teorią statystyczną i w wielu przypadkach zastosowań tej teorii, odpowiednie założenia nie są spełnione. Prowadzi to, oczywiście, do przekłamanych wyników i morza nieprawdziwych prac naukowych.

Wraz z nadejściem komputerów i powszechnej dostępności ogromnej mocy obliczeniowej pojawiły się nowe metody szacowania estymatorów statystyk i ich istotności statystycznej. Metody te, wspólnie nazywane jako *resampling*, nie wymagają znajomości rozkładów prawdopodobieństwa, a założenia gwarantujące ich poprawność są dużo słabsze niż analogiczne założenia w klasycznej statystyce. Wyróżnia się trzy główne kategorie metod resamplingu:

1. bootstrap – oszacowanie dokładności statystyk (mediana, wariancja, percentyl),
2. testy permutacyjne - metoda przeprowadzania testów statystycznych,
3. testowanie krzyżowe - walidacja modeli.

W trakcie referatu opowiem o metodach 1 i 2, ich aspekcie teoretycznym, oraz zaprezentuję przykłady syntetyczne i pochodzące z zastosowań w badaniach.