

## Praca statystyka

- Tworzenie nowych metod analizy danych (jakościowo nowe zbiory danych, nowe kryteria optymalności itp)
- Analiza własności metod statystycznych
  - Teoretyczna (twierdzenia, dowody) – klasyczna statystyka matematyczna
  - Empiryczna – techniki walidacyjne, symulacje komputerowe
- Doradztwo

## Dlaczego warto promować statystykę stosowaną

Małgorzata Bogdan

Instytut Matematyki i Informatyki, Politechnika Wrocławska

Rzeszów, 17/09/2013

## Doradztwo – planowanie eksperymentu i analiza danych

- Zrozumienie problemu praktycznego (np. lokalizacja genów wpływających na zadane cechy)
- Wybór modelu statystycznego/kryterium optymalności (rzadka regresja liniowa/logistyczna, regresja logiczna; kontrola frakcji fałszywych odkryć, maksymalizacja mocy, maksymalizacja zysku, minimalizacja błędu predykcji)
- Planowanie eksperymentu (wybór krzyżówki eksperymentalnej, wstępna selekcja osobników)
- Analiza danych (problem danych brakujących, często istniejące metody wymagają modyfikacji)
- Przełożenie wyników matematycznych na język danej nauki

## Statystyka

- Gałąź matematyki dotycząca metod analizy danych doświadczalnych, planowania eksperymentów losowych oraz podejmowania decyzji w warunkach niepewności przy wykorzystaniu formalizmu matematycznego i metod matematycznych.
- Jerzy Neyman - eksperymenty rolnicze
- Ronald Fisher - genetyka

### Korzyści z rozwoju statystyki/matematyki stosowanej (2)

- Więcej dobrych i przedsiębiorczych studentów na studiach z matematyki – widzą szansę na dobry zawód (konieczność zmiany standardowych programów kształcenia)  
Spośród tych studentów łatwiej wybrać jednostki zdolne do uprawiania matematyki teoretycznej
- Źródło nowych teoretycznych problemów
- Granty i wysoko punktowane publikacje dla Instytutów Matematycznych
- Dodatkowe środki z doradztwa

### Rozwój statystyki w ostatnich latach

- Rozwój technologii pomiarowych i komputerowych
- Duże zbiory danych – konieczne nowe pojęcia i nowe metody (rozwój teorii)  
Klasykne metody analizy generują bardzo dużo fałszywych odkryć
  - testowanie wielokrotne, frakcja fałszywych odkryć
  - compressed sensing – identyfikacja nisko wymiarowych struktur w losowych (zaszumionych) danych wielowymiarowych (analiza obrazów, MRI, X-ray crystallography, analiza danych genetycznych, dopełnianie macierzy)
  - modele hierarchiczne w statystyce bayesowskiej (Markov Chain Monte Carlo) - elastyczne modelowanie wiedzy apriori
  - nieparametryczna statystyka bayesowska
  - nowe kryteria wyboru modelu

### Statystyk stosowany

- Statystyk to nie jest gorszy matematyk
- Pożądane cechy statystyka stosowanego
  - Bardzo dobra znajomość różnych gałęzi matematyki i zaawansowanych metod statystycznych
  - Biegła znajomość narzędzi informatycznych
  - Umiejętność komunikacji ze światem zewnętrznym – umiejętność identyfikacji/określenia problemu matematycznego

### Korzyści z rozwoju statystyki/matematyki stosowanej (1)

- Przekaz dla społeczeństwa - wiele bardzo ważnych problemów praktycznych można rozwiązać jedynie za pomocą zaawansowanych metod matematycznych
  - wczesna diagnoza raka w oparciu o dane „bioinformatyczne” – ekspresja genów, licznosc protein
  - identyfikacja komórek rakowych (np. topologia)
  - planowanie i analiza badań klinicznych (testowanie i wdrażanie nowych leków)
  - indywidualna terapia – identyfikacja osobniczych czynników wpływających na skuteczność terapii

#### Zasady promocji z matematyki stosowanej

- Komisja Zastosowań Matematyki
- Rozważenie poszerzenia kryteriów i sposobu oceny dorobku przy przewodach doktorskich, habilitacyjnych i profesorskich tak by uwzględnić specyfikę wyłaniających się nowych specjalności z matematyki stosowanej
- Rozważenie możliwości powoływania do komisji habilitacyjnych specjalisty z dziedziny, której dotyczy zastosowanie.

#### Zasady promocji (1)

- Kryteria oceny rozpraw doktorskich i habilitacyjnych ze statystyki matematycznej
- Zespół w składzie: prof. Tadeusz Bednarski, prof. Tadeusz Caliński, prof. Jan Mielniczuk (przewodniczący) i prof. Wojciech Niemiro
- Problematyka traktowana w pracach doktorskich i w dorobku habilitacyjnym powinna być ujęta w języku matematycznym.

#### Zasady promocji (2)

W statystyce matematycznej oryginalne rozwiązanie problemu naukowego może obejmować:

- zaproponowanie i zbadanie nowych metod lub podanie nowych własności znanych metod w ogólnych problemach wnioskowania
- budowę lub rozwinięcie istniejących modeli statystycznych w odniesieniu do potencjalnie nowych zagadnień praktycznych
- oryginalne badania empiryczne dla określenia nowych związków pomiędzy istniejącym modelem probabilistycznym i badanym zjawiskiem
- oryginalne badania empiryczne wymagające udoskonaleń istniejących metod analitycznych lub obliczeniowych
- Analiza problemu naukowego i jego rozwiązanie może dokonywać się przy użyciu formalizmu matematycznego, jak i przy wykorzystaniu metod opartych na analizie danych i symulacjach.