



Szczegółowy opis świadczonych usług Centrum Analiz Biostatystycznych i Bioinformatycznych

Szczegółowy opis świadczonych usług

Centrum świadczy szeroki zakres kompleksowych usług z zakresu zaawansowanej analizy biostatystycznej i bioinformatycznej. Usługi oferowane są przede wszystkim naukowcom prowadzącym badania kliniczne, epidemiologiczne, populacyjne, badania molekularne oraz translacyjne. Usługi oferowane przez Centrum oparte są o metody statystyczne i techniki obliczeniowe zawarte na niżej przedstawionej liście:

Planowanie eksperymentów

- Precyzowanie hipotez roboczych i zerowych
- Dobór właściwych metod analizy statystycznej i testów
- Dobór właściwej próby kontrolnej
- Estymacja liczebności próby
- Metody randomizacji

Podstawowe statystyki opisowe

- Statystyki opisowe (średnia, SD, mediana, IQR, zakres, itp.)
- Statystyki opisowe (tabele podsumowujące)
- Podstawowe macierze korelacji
- Tablice wielodzielcze (testy χ^2)
- Tablice częstości
- Testy do oceny normalności rozkładu
- Pole powierzchni pod krzywą (AUC)
- Przedziały ufności

Ocena rozkładu danych

- Test W Shapiro-Wilka
- Test Kolmogorov-Smirnov

Testy t Studenta

- Test t Studenta dla pojedynczej próby
- Test t Studenta dla prób sparowanych
- Test t Studenta dla dwóch prób niezależnych

Testy istotności dla więcej niż 2 prób: analiza wariancji (ANOVA)

- Jednoczynnikowa analiza wariancji (ANOVA)



- Analiza kowariancji (ANCOVA)
- Wielowymiarowa analiza wariancji (MANOVA)
- Analiza wariancja w układach z nierówną liczbą powtórzeń w grupach
- Ogólny model liniowy (GLM)
- Analiza wariancji z powtarzаныmi pomiarami
- Pole powierzchni pod krzywą (AUC)
- Testy istotności dla porównań wielokrotnych *post-hoc* (test *t* Studenta z poprawką Bonferroni'ego, test LSD Fishera, test Scheffe'go, test Tukey'a, test Duncana, itp.)
- Testy istotności dla porównań zaplanowanych (analiza kontrastów, test Dunnetta, itp.)
- Testy umożliwiające sprawdzenie założeń analizy (test *C* Cochрана, test Hartley'a, test Bartletta, test Levene'a)

Ocena zgodności metod

- Wykres i analiza Bland-Altmana
- Testy oparte na statystyce χ^2 i *z*

Testy istotności dla proporcji oparte o statystyki χ^2 i *z*

Analiza korelacji

- Korelacje Pearsona, Spearmana, *Tau* Kendalla, Gamma
- Korelacje cząstkowe w ramach regresji wielorakiej
- Macierze korelacji
- Korelacje funkcji kanonicznych

Analiza regresji

- Regresja liniowa
- Regresja wieloraka i korelacje cząstkowe
- Analiza kowariancji (ANCOVA)
- Ogólny model liniowy (GLM)
- Regresja logistyczna
- Regresja Coxa
- Regresja probitowa
- Regresja nieliniowa

Metody dopasowania krzywej

- Metody dopasowania krzywej rozkładu (normal, exponential, gamma, log-normal, χ^2 , Weibull, binomial, Poisson, Bernoulli, itp.)
- Regresja wielomianowa
- Regresja nieliniowa

Analiza tablic wielodzzielczych

- Tablice wielodzzielcze (testy χ^2)
- Tablice częstotliwości
- Test Q Cochрана
- Modele analizy log-liniowej
- Test Mantel-Haenszela



- Testy chi^2 dla trendu

Metody nieparametryczne

- Test U Manna-Whitney'a
- Test H Kruskala-Wallisa
- Test Friedmana
- Analiza skumulowanych częstości
- Korelacje Spearmana, Tau Kendalla, Gamma
- Tablice wielodzielcze (Chi-Square Test)
- Test dla prób zależnych McNemara
- Test Q Cochra
- Testy Wilcoxon
- Testy zgodności chi^2
- Krzywe Kaplana-Meiera (testy log-rank)
- Analiza ROC i pola powierzchni pod krzywą (AUC)

Analiza krzywych ROC

- Analiza pojedynczej krzywej ROC i wyznaczenie wartości cut-off
- Porównanie dwóch krzywych ROC
- Pole powierzchni pod krzywą (AUC)

Analizy wielowymiarowe

- Regresja logistyczna
- Ogólne modele liniowe
- Uogólnione modele liniowe i nieliniowe
- Analiza czynnikowa
- Analiza korelacji funkcji kanonicznych
- Analiza funkcji dyskryminacyjnej
- Analiza składowych głównych
- Analiza skupień
- Wielowymiarowa analiza wariancji (MANOVA)
- Analiza korespondencji
- Analiza log-liniowa tablic wielodzielczych
- Analiza przeżycia
- Regresja proporcjonalnego hazardu Coxa
- Estymacja i regresja nieliniowa



Analiza skupień

- Metody rozmytej analizy skupień (fuzzy clustering)
- Metody hierarchiczne/Dendrogramy
- Grupowanie metodą k-średnich (k-means clustering)
- Podwójne dendrogramy (heatmaps)

Analiza przeżycia

- Krzywe skumulowanej częstości występowania
- Krzywe Kaplana-Meiera (testy logrank)
- Regresja Coxa

Meta-analizy

- Meta-analiza dla wartości RR, OR
- Meta-analiza dla średnich
- Meta-analiza dla proporcji

Techniki eksploracyjne, data mining i machine learning

- Drzewa klasyfikacyjne CART i CHAID
- Klasyfikacja metodą lasu losowego (Random Forest)
- Regresja MARSplines
- Techniki z wykorzystaniem maszyn wektorów nośnych (Support Vector Machines)
- Sieci neuronowe

Procedury graficzne

- Wykresy kolumnowe 2D i 3D
 - dla jednego czynnika
 - dla wielu czynników
- Wykresy zgodności Bland-Altmana
- Wykresy pudełkowe (skrzynka-wąsy)
- Histogramy
 - histogramy okrągłe
 - histogramy kolumnowe
 - histogramy kołowe
- Wykresy konturowe
- Wykres powierzchni 3D
- Dendrogramy
 - dendrogramy hierarchiczne
 - podwójne dendrogramy (heatmaps)
- Wykresy punktowe



- Wykresy słupkowe
- Wykresy Forest Plot dla meta-analiz
- Heatmapy
 - podwójne dendrogramy
 - macierze korelacji
 - wykresy powierzchni 3D
 - wykresy konturowy
- Krzywe Kaplan-Meiera
- Krzywe ROC
- Wykresy rozproszenia danych
 - wykresy rozproszenia 2D i 3D
 - wykresy rozproszenia ze słupkami błędów
- Wykresy 3D
 - wykresy rozproszenia danych 3D
 - wykresy powierzchni 3D
 - wykresy kolumnowe 3D
 - wykresy analizy korespondencji

Opracowanie danych

- Import danych
- Export danych
- Filtrowanie i czyszczenie danych
- Transformacja danych
- Normalizacja danych
- Randomizacja
- Generowanie symulowanych zbiorów danych
- Łączenie zbiorów danych

Przygotowanie raportów

- Streszczenie (zwięzłe przedstawienie celu analizy, zastosowanych metod, najważniejszych wyników i wniosków)
- Podsumowanie analizowanych problemów badawczych oraz testowanych hipotez
- Szczegółowy opis zastosowanej procedury analitycznej oraz metod
- Szczegółowe zestawienie i omówienie uzyskanych wyników analizy
- Szczegółowe omówienie wniosków wyciągniętych z analizy statystycznej
- Zwrócenie uwagi na silne i słabe strony doświadczenia oraz przeprowadzonej analizy
- Sugestie i propozycje na przyszłość



Analizy bioinformatyczne

- Alignment sekwencji nukleotydowych/białkowych
- Przeszukiwanie biologicznych baz danych
- Annotacja regionów sekwencji
- Annotacja funkcjonalna sekwencji
- Analizy filogenetyczne
- Umieszczanie sekwencji/bibliotek odczytów z sekwencjonowania w bazach danych
- Identyfikacja specyficznych motywów sekwencji
- Składanie de novo genomów organellowych oraz genomów bakterii/grzybów
- Identyfikacja wariantów splicingowych
- Analiza zmienności genetycznej
- Profilowanie ekspresji genów
- Analiza zróżnicowanej ekspresji genów (bulk RNA-seq, scRNA-seq, mikromacierze)
- Analiza zróżnicowanej ekspresji oraz profilowanie miRNA
- Analiza profilu mikrobiomu w oparciu o sekwencjonowanie 16S rRNA oraz WGS
- Identyfikacja metylacji w odczytach nanoporowych (ONT)
- Analiza miejsc wiązania czynników transkrypcyjnych i histonów (ChIP-seq)
- Pomoc w zaplanowaniu eksperymentu opartego o sekwencjonowanie nowej generacji
- Wizualizacja danych

Konsultacje statystyczne i bioinformatyczne