Position Announcement: Applied Statistician

Dept: UC DAVIS GENOME CENTER - 011017
Position: Project Scientist
HEERA/Union Representation: This position is covered by a collective bargaining unit

IMMEDIATE SUPERVISOR

Supervisor Name: Oliver Fiehn
Supervisor Payroll Title: Professor
Supervisor Phone Number: 4-8258

POSITION DETAILS

The West Coast Metabolomics Center seeks an experienced, applied statistician to collaborate with biochemists, analytical chemists, informaticists and NIH-funded principal investigators to design, analyze and interpret metabolomic studies in a range of biomedical and clinical projects. The applied statistician will provide expertise, advice and perform statistical analyses for the Center and the Center's collaborators.

The candidate will have knowledge and experience working with large complex datasets for multivariate and univariate tests, as well as demonstrated expertise with respect to factors that impact and bias study designs and which possibly could confound data sets. The successful candidate will have experience in working with teams from different scientific backgrounds to explore data-driven hypotheses as well as hypothesis-driven studies. The applied statistician will perform various types of corrections for multiple testing, including non-parametric methods. The candidate must be able to calculate univariate and multivariate receiver-operator curves for estimating impact of variables on disease phenotypes. The applied statistician will perform power calculations for planned studies involving univariate and multivariate datasets, including non-ideal data structures. In order to provide fast services for metabolomic studies, the statistician will develop the use of a suite of default multivariate methods such as PCA, PLS, OPLS-DA, ICA or related methods in addition to univariate tests to provide clients with preliminary data visualizations and statistical outputs. The incumbent will test for statistical significance in a variety of ways, testing for model overfitting and perform biomarker calculations using an optimal number of variables, including predictive model optimization, validation, and feature selection.

The ability to develop, implement and perform novel types of statistical tests and to derive appropriate models in non-textbook situations is essential ,including scientific programming and algorithm development. The applied statistician will be able to analyze time-dependent data as well as data sets from multifactorial studies. The statistician has knowledge of the effect of normalization and transformation treatments to the structure of data sets and is able to explain the choice of normalization tools to statistically untrained scientists.

Multivariate data in disease and cellular biology are not independent. The candidate will go beyond the assumption of mutual independence of variables and engage, test, develop and help implement further statistical tools that take the correlative nature of biological data into consideration including application of partial correlations, including conditional independence and construction of graphical Gaussian Markov models.

The statistician will have expertise in gene enrichment analysis or related tools such as pathway enrichment or cluster enrichment tests. The incumbent will help the West Coast Metabolomics Center to develop such tools as innovative ways to test for significant dysregulation of metabolites in biological studies.

The statistician will participate in team meetings and explain findings to groups of Center scientists and Center collaborators. The candidate must have demonstrable ability to communicate well in these complex situations. The applied statistician is expected to have experience with multiple computational and visualization tools, including but not limited to R, Matlab, or Statistica. The successful candidate will hold at least a Master's degree in statistics or equivalent with multiple years of practice analyzing biological or clinical data sets.

QUALIFICATIONS
Minimum Qualifications:

- Master's degree in statistics or equivalent. PhD preferred.
- Significant experience with the applied statistics of large datasets.
- Demonstrated ability to use a variety of software for statistical testing, including but not limited to R, Matlab and Statistica.
- Significant experience in multiple testing problems, non-normally distributed datasets, non-parametric statistics, datasets with many internally correlated measures.
- Significant experience with clinical datasets, ROC curves, power calculations.
- Demonstrated ability to generate data visualizations and graphical outputs.
- Experience with development of new metrics, normalizations and data transformations.
- Demonstrated experience reading and applying methods from the statistical literature.
- Experience summarizing and presenting data in writing and to small and large groups.
- Ability to skillfully communicate statistical results to researchers and collaborators

Preferred Qualifications:

- General familiarity with mass spectrometry, genomics, proteomics and related fields.
- Analysis and validation of biomarkers.
- Experience integrating multiple datasets from different data streams: eg. genomics, proteomics, metabolomics or related measurements.
- Experience with data networks and analysis and/or cluster analysis.