

# Brainstorm: Statistics

## Distributions

Normal  
Gamma  
Bernouli  
Binomial  
t-distribution  
Weibull  
Negative Binomial  
Cauchy  
Poisson  
Chi-squared  
Beta  
Laplace  
Geometric  
Exponential  
Uniform  
F  
Normal-chi  
log-normal  
Pareto

# Statistical Models

(General) Linear Model  
Random Forest  
Mixed linear models  
Lasso / Ridge regression / Elastic Net  
Polynomial models  
Non-linear models  
L/QDA - linear/quadratic discriminant analysis  
Logistic regression  
Bayesian networks  
SVM - support vector machines  
Hierarchical clustering  
Generalized linear models  
Hidden Markov models  
Neural networks  
Kriging  
Gaussian mixture models  
Autoregressive (ARIMA)  
Generalized additive models (GAMs)  
Box-Cox models  
Weibull regression  
Cox proportional hazards model  
Monte Carlo simulations  
Accelerated failure time models  
ANOVA (analysis of variance)  
Non-parametric regression  
Kernel regression  
Nearest neighbour  
Spline functions  
t-SNE / UMAP  
Markov random fields

## Methods of estimation

Maximum likelihood estimation (MLE)

Bayesian inference

Method of moments

Expectation-Maximization (EM)

rank-based methods

OLS - ordinary least squares

Bootstrap estimation

SSR - sum of squared residuals

Markov chain Monte Carlo (MCMC)

Generalized/iteratively (re)weighted least squares

Generalized estimating equations (GEEs)

# Hypothesis testing

Student's t-test

ANOVA

Wald test

Fisher exact test

Chi-squared test (of independence)

ANCOVA

F test

Score test

Wilcoxon (signed rank)

Kolmogorov-Smirnov test

Kruskal-Wallis

Tukey/Dunn's/etc post hoc tests

Likelihood Ratio test

Mann-Whitney

Spearman rank test

log-rank test

Shapiro-Wilke / Anderson-Darling

Bartlett's test

## Technologies (in molecular biology)

High throughput sequencing  
Polymerase chain reaction (PCR)  
ELISA  
Microarray  
Northern/Southern/Eastern/Western blots  
Single cell sequencing  
FACS (fluorescence activated cell sorting) / flow cytometry  
CITE-seq (Ab-seq)  
Mass cytometry  
(Chromatin Immunoprecipitation) ChIP-seq  
FISH - fluorescence in situ hybridization  
FRET - fluorescence resonance energy transfer  
HiC  
RNA-seq (RNA sequencing)  
Bisulphite sequencing (BS-seq)  
HPLC (liquid chromatography)  
CRISPR  
Mass spectrometry  
NMR  
Reverse transcription  
DSTORM  
microfluidics  
Gel electrophoresis  
Magnetic tweezers  
Nanopore sequencing  
Long-read sequencing (Nanopore, Pac Bio)  
Nanodroplets  
Microscopy  
SPR - surface plasma resonance

# Applications

Forensics

Biodiversity and conservation

Finding disease-associated mutations

GWAS - genome-wide association studies

Vaccines

Understand tissue heterogeneity

Chromatin accessibility / states - treatment/lifestyle on epigenetics

Determine factors important in development

Cell type identification

Basic physiological processes of a cell

Gene -> protein expression -> targets of a drug

Diagnostics / biomarker

Molecular underpinnings of cancer

Pathway analysis

RNA localization

Cell plasticity

Foundations for metabolic engineering

Linking technology -> applications -> statistics

| Technology     | Applications                 | Statistics               |
|----------------|------------------------------|--------------------------|
| Microarray     | Differential Gene Expression | Linear models            |
| RNA Sequencing | Differential Gene Expression | Negative Binomial models |
|                |                              |                          |