

TIGP Bio 2022 Fall Syllabus
Fundamental Statistical Methods in Bioinformatics (S1)

For the latest syllabus, please visit the BP website: <https://tigbp.iis.sinica.edu.tw>

Place: Room 107, New Building of the Institute of Information Science, Academia Sinica
Time: Thursday 9:00-12:00
Chair: Dr. Shinsheng Yuan (syuan@stat.sinica.edu.tw; shinshengyuan@gmail.com)
Outline: This course covers the fundamentals of statistics and basic tools for bioinformatics analysis. In the first part students will learn basic statistical concepts and methods, including probability, random variables and distributions, parameter estimation, hypothesis testing, regression analysis, and categorical data analysis. In the second part several commonly used methods in bioinformatics will be introduced, including statistical meta analysis, survival analysis, clustering, classification, and nonparametric statistics.
Textbook: Fundamentals of Biostatistics (author: Bernard Rosner), Cengage Learning.
Reference book: Pattern Recognition (author: Richard O. Duda, Peter E. Hart, and David G. Stork), Wiley.
TA: Tzu-Hsiang Lin **Email:** tzuhsiang@gate.sinica.edu.tw
Office hours: Tuesdays, 18:30~20:30 pm **Office location:** Room A622, Agricultural Biotechnology Research Centers, Academia Sinica
Grades: Midterm exam 50%. Final exam 50%.

【For Non-BP student】

For Non-BP student to register/sit-in any BP course, it is required to gain course chair's permission:

- (1) Basic Enrollment Information form (<http://bit.ly/3g22mfl>)
(2) TIGP-BP Course Registration Consent Form (<http://bit.ly/3ubl2Bs>)

※ Deadline: **the 4th week** of each semester.

※ Signature of the course chair should be collected before submission. Incomplete form will not be accepted.

※ Course grade will NOT be given (even class enrollment is completed at school) if fail to follow the above procedures.

Week	Date	Topics/Brief Description	Sub-topics/Detail Descriptions	Lecturers	Evaluation Method
1	2022/9/8	Descriptive Statistics, Genomic Data Analysis	(1) Introduction to statistics (2) Descriptive statistics (3) Fundamental of molecular biology (4) Genomic data analysis	Dr. Chen-Hsiang Yeang	HW with a specified deadline: Oct. 27th
2	2022/9/15	Probability	(1) Applications in statistical genetics (2) Combinatorial analysis (3) Axioms of probabilities (4) Conditional probability and independence (5) Random variable and distribution function	Dr. Hsin-Chou Yang	HW with a specified deadline: Oct. 16th
3	2022/9/22	Discrete Distributions and Contingency Tables	(1) An application in pharmacogenetic study (2) Discrete/continuous/mixed distributions (3) Joint/marginal/conditional distributions (4) Special discrete distributions (5) Introduction to contingency table	Dr. Hsin-Chou Yang	
4	2022/9/29	Continuous distributions and basic statistics	(1) Continuous random variable (2) Expectation (3) Basic statistics (4) Limit theorems (optional)	Dr. Hsin-Chou Yang	
5	2022/10/6	Parameter Estimation and Confidence Interval	(1) Unbiasedness (2) Point estimation (substitution principles, least square estimate, maximum likelihood estimate) (3) Interval estimation	Dr. Shin-Sheng Yuan	in-class exam on midterm exam date
6	2022/10/13	Hypothesis Testing, P-value and False Discovery Rate	(1) Hypothesis testing (2) Applications in cancer researches (3) Type I & type II errors and p-value (4) One-sample and two-sample z-tests (5) One-sample, two-sample, and paired t-tests (6) Bonferroni adjustment, false discovery rate, and q value	Dr. Grace S. Shieh	in-class exam on midterm exam date
7	2022/10/20	Review Week (No class)			
8	2022/10/27	Midterm Exam	--	--	--
9	2022/11/3	Regression Analysis	(1) Applications (2) Simple linear regression & inference (3) Diagnostic and remedial measures (4) Matrix approach to simple linear regression (5) Multiple linear regression (6) Building the regression model	Dr. Grace S. Shieh	in-class exam on final exam date
10	2022/11/10	Survival Data Analysis	(1) Mantel-Haenszel test (2) Survival and hazard functions (3) Kaplan Meier estimate (4) Log-rank test (5) Proportional-hazards model (6) Lung cancer study	Dr. Hsuan-Yu Chen	HW with a specified deadline: ▶ Dec. 29th
11	2022/11/17	Logistic Regression and Statistical Meta Analysis	(1) Logistic regression (2) Meta analysis (effect size, precision, study weights, summary effect, heterogeneity, fixed-effect model, random-effect model, software)	Dr. Shin-Sheng Yuan	HW with a specified deadline: ▶ Dec. 26th
12	2022/11/24	Clustering	(1) Clustering by geometry (K-means, EM algorithm, hierarchical clustering, self-organizing map, principal component analysis, independent component analysis) (2) Clustering on graphs (Basic concepts, max flow – min cut, normal cuts, spectral clustering, and community detection) (3) Advanced topics (Chinese restaurant process and affinity propagation)	Dr. Chen-Hsiang Yeang	in-class exam on final exam date

13	2022/12/8	Nonparametric Statistics (I)	(1) Bootstrap (2) One-sample sign test (3) One-sample Wilcoxon signed-rank test (4) Wilcoxon rank-sum test (Mann-Whitney U test) (5) Sign test for paired data (6) Wilcoxon signed-rank test for paired data	Dr. Wei-Chung Liu	
14	2022/12/14 (Wed) 9:00-12:00	Nonparametric Statistics (II)	(1) Kruskal-Wallis test (2) Randomization/permutation test for two-way ANOVA (3) The product-moment correlation coefficient (4) Spearman rank correlation (5) Kendall's coefficient of rank correlation	Dr. Wei-Chung Liu	HW with a specified deadline: ▶ Jan. 1st
15	2022/12/15	Review Week (No class)	--	--	--
16	2022/12/22	Final Exam	--	--	--