《统计推断》课程教学大纲

一、课程基本信息

英文名称	Statistical Inference	**课程代码**	FIAI1027
课程性质	专业选修课程	**授课对象**	国际金融专业
学　　分	3.0	**学　　时**	54
主讲教师	外教	**修订日期**	2020.3
指定教材	R. Lyman Ott. An Introduction to Statistical Methods and Data Analysis 6th Edition. Cengage Learning; 6 edition		

二、课程描述

Statistical Inference focuses on making inferences about population quantities from sample data via hypothesis testing and confidence intervals. Statistical Inference applies statistical methods in a business context in order to address business related questions and help make evidence based decisions. In Statistical Inference, you will learn to apply commonly used statistical methods in business contexts and how to interpret analyses performed by others.

"Statistical inference refers to inferring conclusions or making decision based on the data. Two of the most important topics will be covered in this course, namely, parameter estimation and hypothesis testing. The methods for statistical inference entails our understanding of the randomness. Data is assumed to be random and is described by a probability distribution function. Traditionally, the probability distributions are assumed to be following a model that is characterized by some parameters. The choice of the model reflects our belief on the properties of the data. In reality, we can never observe the values of the parameters, however, the data contains in some sense the information about them. Our objectives include estimating the unknown parameters, testing if the data fits a specified model, comparing the parameters of two sets of data, etc. This course introduces some formal procedures for constructing the estimators of the unknown parameters and the criterion of accepting / rejecting an hypothesis. A review of probability theory, sampling distribution theory and large sample theory will be given to the students before presenting the methods of estimation and hypothesis testing. Students completing the course will be able to:

1. Create and interpret scatter plots and histograms.
2. Understand the difference between probability and likelihood functions, and find

the maximum likelihood estimate for a model parameter.

3. Construct estimates and predictions using the posterior distribution.

4. Find credible intervals for parameter estimates.

5. Use null hypothesis significance testing (NHST) to test the significance of results, and understand and compute the p-value for these tests.

6. Use specicific significance tests including, z-test t-test (one and two sample), chi-squared test.

7. Find confidence intervals for parameter estimates.

8. Compute and interpret simple linear regression between two variables.

9. Set up a least squares fit of data to a model

三、课程大纲

Topics
Estimation, confidence intervals and hypotheses tests. Simple regression and correlation. Multiple regression; T and F tests.

Topic 1: Introductory Statistical Language
Understand basic principles of statistical inference & Become an informed consumer of statistical information.

Topic 2: Organizing and Visualizing Data, Numerical Descriptive Techniques
Create and interpret scatter plots and histograms. 1. Single variable 1. Single variable – Part I (Basic)
1.1. How to calculate and use the measures of location
1.2. How to calculate and use the measures of variability
2. Single variable – Part II (Application)
2.1. Understand what the measures of location (e.g., mean, median, mode) tell us about distribution shape - Discuss its use in manipulating simulated experiments
2.2. How to detect outliers using z-score and empirical rule
2.3. How to use Box plot to explore data
2.4. How to calculate weighted mean
2.5. How to calculate mean and variance for grouped data
3. Two variables
3.1. How to calculate and use the measures of association - Covariance, Correlation coefficient
1.1. How to use Tables and Graphs to summarize data
1.1.1. Qualitative data
1.1.2. Quantitative data
1.2. How to use Stem-and-Leaf display to explore data
2. Two variables

2.1. How to identify and understand potential relationship between variables
2.1.1. Using Crosstabulation – Power of Simpson's paradox when present
2.1.2. Using Scatter Diagram and Trendline
Understand the difference between probability and likelihood functions, and find the maximum likelihood estimate for a model parameter.
Do Bayesian updating with discrete priors to compute posterior distributions and posterior odds.
Do Bayesian updating with continuous priors.
Construct estimates and predictions using the posterior distribution.
Find credible intervals for parameter estimates.

Topic 3: Probability
Learn the language and core concepts of probability theory: Counting Random variables, distributions, quantiles, mean variance Conditional probability, Bayes' theorem, base rate fallacy Joint distributions, covariance, correlation, independence Central limit theorem 1. Introduction to probability
1.1. Understand experiments, outcomes, sample space.
1.2. How to assign probabilities to outcomes
1.3. Understand events and how to assign probabilities to them
2. Basic relationships of probability
2.1. Understand and know how to compute:
- complementary event,
- union of two events,
- intersection of two events,
- mutually exclusive events.
3. Conditional probability
3.1. Understand independent events
3.2. Understand multiplication law
4. Bayes' Theorem
4.1. Tabular approach

Topic 4: Random Variables and Discrete Probability Distributions
1. Understand random variables and probability distributions.
1.1. Distinguish discrete and continuous random variables.
2. Able to compute Expected value and Variance of discrete random variable.
3. Understand:
3.1. Discrete uniform distribution
3.2. Binomial distribution
3.3. Poisson distribution

Topic 5: Continuous Probability Distributions
1. Understand continuous probability distributions
2. Understand Uniform distribution
3. Understand Normal distribution

3.1. Understand Standard normal distribution
3.2. Understand Normal approximation of binomial distribution
4. Understand Exponential distribution
4.1. Understand relationship between Poisson and Exponetial distribution.

Topic 6: Sampling Distribution and Central Limit Theorem & Data Collection and Sampling Methods
1. Understand Simple Random Sampling
2. Understand Point Estimation and be able to compute point estimates
3. Understand properties of Point Estimators
4. Understand other Sampling Methods

Topic7: Confidence Interval Estimation and Statistical Inference
1. Compute confidence interval for Population Mean when s is known
2. Compute confidence interval for Population Mean when s is unknown
3. Compute appropriate sample size for given confidence level.
4. Compute confidence interval for population proportion.

Topic 8: Hypothesis Testing
1. Understand how to develop Null and Alternative Hypotheses
2. Understand Type I and Type II Errors
3. Able to do hypothesis test about population mean when σ is known
4. Able to do hypothesis test about population mean when σ is unknown

Topic 9: Two Sample Tests & Linear Regression Analysis
1. Define the explanatory variable as the independent variable (predictor), and the response variable as the dependent variable (predicted).
2. Plot the explanatory variable (x) on the x-axis and the response variable (y) on the y-axis, and fit a linear regression model $y = \beta0 + \beta1x$, where β0 is the intercept, and β1 is the slope. - Note that the point estimates (estimated from observed data) for β0 and β1 are b0 and b1, respectively.
3. When describing the association between two numerical variables, evaluate - direction: positive (x ↑, y ↑), negative (x ↓, y ↑) - form: linear or not - strength: determined by the scatter around the underlying relationship
4. Define correlation as the linear association between two numerical variables.
A. Understand how regression analysis can be used to develop an equation that estimates mathematically how two variables are related
B. Understand the differences between the regression model, the regression equation, and the estimated regression equation.
C. Know how to fit an estimated regression equation to a set of sample data based upon the least-squares method.
D. Be able to determine how good a fit is provided by the estimated regression equation and compute the sample correlation coefficient from the regression analysis output.

E. Understand the assumptions necessary for statistical inference and be able to test for a significant relationship.
F. Know how to develop confidence interval estimates of y given a specific value of x in both the case of a mean value of y and an individual value of y.
G. Learn how to use a residual plot to make a judgement as to the validity of the regression assumptions, recognize outliers, and identify influential observations.

四、教学进度

周次	教学内容 **Teaching content**	学时分配 **Class hour**	目的要求 **Purpose requirements**
1	Introductory Statistical Language	3	Understand basic principles of statistical inference & Become an informed consumer of statistical information.
2	Organizing and Visualizing Data	3	Create and interpret scatter plots and histograms. 1. Single variable 1. Single variable – Part I (Basic) 1.1. How to calculate and use the measures of location 1.2. How to calculate and use the measures of variability 2. Single variable – Part II (Application) 2.1. Understand what the measures of location (e.g., mean, median, mode) tell us about distribution shape - Discuss its use in manipulating simulated experiments 2.2. How to detect outliers using z-score and empirical rule
3	Numerical Descriptive Techniques I	3	

4	Numerical Descriptive Techniques II	3	2.3. How to use Box plot to explore data 2.4. How to calculate weighted mean 2.5. How to calculate mean and variance for grouped data 3. Two variables 3.1. How to calculate and use the measures of association - Covariance, Correlation coefficient 1.1. How to use Tables and Graphs to summarize data 1.1.1. Qualitative data 1.1.2. Quantitative data 1.2. How to use Stem-and-Leaf display to explore data 2. Two variables
5	Numerical Descriptive Techniques III	3	2.1. How to identify and understand potential relationship between variables 2.1.1. Using Crosstabulation – Power of Simpson's paradox when present 2.1.2. Using Scatter Diagram and Trendline Understand the difference between probability and likelihood functions, and find the maximum likelihood estimate for a model parameter. Do Bayesian updating with discrete priors to compute posterior distributions and posterior odds. Do Bayesian updating with continuous priors. Construct estimates and predictions using the posterior distribution. Find credible intervals for parameter estimates.

6	Probability	3	Learn the language and core concepts of probability theory: Counting Random variables, distributions, quantiles, mean variance Conditional probability, Bayes' theorem, base rate fallacy Joint distributions, covariance, correlation, independence Central limit theorem 1. Introduction to probability 1.1. Understand experiments, outcomes, sample space. 1.2. How to assign probabilities to outcomes 1.3. Understand events and how to assign probabilities to them 2. Basic relationships of probability 2.1. Understand and know how to compute: - complementary event, - union of two events, - intersection of two events, - mutually exclusive events. 3. Conditional probability 3.1. Understand independent events 3.2. Understand multiplication law 4. Bayes' Theorem 4.1. Tabular approach
7	Random Variables and Discrete Probability Distributions		1. Understand random variables and probability distributions. 1.1. Distinguish discrete and continuous random variables. 2. Able to compute Expected value and Variance of discrete random variable. 3. Understand: 3.1. Discrete uniform distribution 3.2. Binomial distribution 3.3. Poisson distribution

8	Continuous Probability Distributions	3	1. Understand continuous probability distributions 2. Understand Uniform distribution 3. Understand Normal distribution 3.1. Understand Standard normal distribution 3.2. Understand Normal approximation of binomial distribution 4. Understand Exponential distribution 4.1. Understand relationship between Poisson and Exponetial distribution.
9	Sampling Distribution and Central Limit Theorem	3	1. Understand Simple Random Sampling 2. Understand Point Estimation and be able to compute point estimates 3. Understand properties of Point Estimators 4. Understand other Sampling Methods
10	Midterm Exam	3	
11	Data Collection and Sampling Methods	3	
12	Confidence Interval Estimation and Statistical Inference	3	1. Compute confidence interval for Population Mean when s is known 2. Compute confidence interval for Population Mean when s is unknown 3. Compute appropriate sample size for given confidence level. 4. Compute confidence interval for population proportion.
13	Hypothesis Testing	3	1. Understand how to develop Null and Alternative Hypotheses 2. Understand Type I and Type II Errors 3. Able to do hypothesis test about population mean when σ is known 4. Able to do hypothesis test about population mean when σ is unknown
14	Hypothesis Testing	3	
15	Two Sample Tests	3	1. Define the explanatory variable as the independent variable (predictor), and the response variable as the dependent variable (predicted). 2. Plot the explanatory variable (x) on the x-axis and the response variable (y) on the y-axis, and fit a linear regression model y = β0 + β1x, where β0 is the intercept, and β1 is the slope. - Note that the point estimates (estimated from observed data) for β0 and β1 are b0 and b1, respectively. 3. When describing the association between two numerical variables, evaluate - direction: positive (x ↑, y ↑), negative (x ↓,

16	Linear Regression Analysis	3	y ↑) - form: linear or not - strength: determined by the scatter around the underlying relationship 4. Define correlation as the linear association between two numerical variables. A. Understand how regression analysis can be used to develop an equation that estimates mathematically how two variables are related B. Understand the differences between the regression model, the regression equation, and the estimated regression equation. C. Know how to fit an estimated regression equation to a set of sample data based upon the least-squares method. D. Be able to determine how good a fit is provided by the estimated regression equation and compute the sample correlation coefficient from the regression analysis output. E. Understand the assumptions necessary for statistical inference and be able to test for a significant relationship. F. Know how to develop confidence interval estimates of y given a specific value of x in both the case of a mean value of y and an individual value of y. G. Learn how to use a residual plot to make a judgement as to the validity of the regression assumptions, recognize outliers, and identify influential observations.
17	Review	3	
18	**Final Exam**	3	

五、考核方式及评定方法

- Attendance and Participation 20%
- Mid-Term 30%
- Final Exam 30%

课程目标	评分标准				
	90-100	**80-89**	**70-79**	**60-69**	**＜60**
	优	**良**	**中**	**合格**	**不合格**
	A	**B**	**C**	**D**	**F**

六、参考书目

Statistics for Business and Economics, 8th edition, Paul Newbold, William L. Carlson, Betty M. Thorne, Pearson.