STAT 1000
Basic Statistical Analysis I

Course Content

Unit 1 – Examining Distributions

- types of variables: quantitative, categorical, continuous, nominal, ordinal
- graphs for categorical variables: bar charts, pie charts
- graphs for quantitative variables: stemplots, histograms
- examining distributions, dealing with outliers
- time plots
- describing distributions with numbers: mean, weighted mean, median, quartiles, percentiles, interquartile range, range, variance and standard deviation
- five-number summary and boxplots
- the $1.5 \times$ IQR rule for suspected outliers, modified boxplots
- resistant measures
- Introduction to `JMP` — a Statistical Computer Package

Unit 2 – Scatterplots, Correlation and Regression

- association, response variable, explanatory variable
- examining scatterplots
- adding categorical variables to scatterplots
- categorical explanatory variables
- correlation, properties of correlation
- least-squares criterion and least-squares regression line
- r^2
- residuals, outliers, influential observations
- cautions about correlation and regression
- association vs. causation, lurking variables
- extrapolation, causation, confounding, establishing causation

Unit 3– Sampling Design

- populations and samples
- voluntary response sample
- simple random sample
- random digits table
- probability sample, stratified random sample, multistage sample

- undercoverage, nonresponse, response bias

Unit 4 – Design of Experiments

- observations vs. experiment
- experimental units, subjects, treatments
- factors, levels
- placebo effect, control group, bias
- randomization, principles of experimental design
- randomized comparative experiments
- randomized block design, matched pairs design
- ethics

Unit 5 – Density Curves and Normal Distributions

- continuous random variables, density curves
- normal distributions
- 68–95–99.7 rule
- standardizing observations (z-scores)
- normal distribution calculations
- use of normal quantile plots

Unit 6 – Randomness and Probability

- randomness, the language of probability
- probability models, sample space, events, unions, intersections
- some probability rules, independence, general addition rule
- discrete random variables
- binomial setting and binomial distribution

Unit 7 – Sampling Distributions

- sampling distribution of a sample mean
- bias and variability
- Central Limit Theorem
- sampling distributions for proportions

Unit 8 – Confidence Intervals for a Population Mean (σ known)

- margin of error
- effect of sample size, confidence level, standard deviation
- effect of population size
- assumptions

- choosing the sample size

Unit 9 – Tests of Significance

- hypotheses, test statistic, P-value, statistical significance
- tests for a population mean (σ known)
- two-sided tests and confidence intervals

Unit 10 – Inference for One or Two Population Means (σ unknown)

- one-sample t procedures — confidence intervals and tests
- robustness of the t procedures

Unit 11 – Inference for a Population Proportion

- sampling distribution of the sample proportion
- confidence intervals and tests
- choosing the sample size

* * * * * * * * * * * * * *

Some Additional Notes

Where there are any discrepancies between the way topics are covered in the course notes and in the reference books, please refer to the notes. Moreover, when there are notations or formulas that differ from the reference books to the notes, please refer to the course notes.

Last updated April 27, 2016.