INTRODUCTORY STATISTICS
Transfer Credit Proposal Draft 4

Split between **core** and **elective** topics
as determined at Joint Articulation Meeting in Calgary, May 2017

SUGGESTED OPTIONS:

- The order of the presentation of the topics is at the discretion of the instructor.
- Choice of statistical software, spreadsheet such as Excel, or calculator
- Both the *p*-Value and Critical-value approaches to hypothesis testing
- With or without use of distribution tables

CORE (75 – 80%)		**ELECTIVE (20 – 25%)**
➢ **Descriptive Statistics**		
Sample Surveys, Sampling Techniques, and Design of Experiments – internal and external (both primary and secondary) sources of data, survey, observational study, experiment, treatment group, control group, blocking, double-blind experiment, random and nonrandom samples, convenience sample, judgment sample, quota sample, sampling error, nonsampling error, selection error, voluntary response error or bias, simple random sampling, systematic random sampling, stratified random sampling, cluster sampling		
Statistics, Population, and Sample – descriptive and inferential statistics, census, sample survey, representative sample, random sample, sampling with replacement, sampling without replacement, element, variable, observation, measurement, data set		
Types of Variables and the Nature of Statistical Data – quantitative, discrete, continuous, qualitative, categorical, parameter, statistic, summation notation		**cross-section, time series**
Organizing and Graphing Qualitative Data – frequency, relative frequency and percentage distributions, bar charts, pie charts		

Organizing and Graphing Quantitative Data – frequency, relative frequency and percentage distributions, classes of data, class width, class midpoint, histograms, frequency polygons, stem-and-leaf plots, pie charts		**class limits and class boundaries of histograms, cumulative frequency distributions, cumulative relative frequency and cumulative percentage distributions, ogives, dotplots**
Measures of Central Tendency for Ungrouped Data – mean, median, mode		
Measures of Dispersion for Ungrouped Data – range, variance, standard deviation		
		Mean, Variance, and Standard Deviation for Grouped Data – for both population and sample data
Use of Standard Deviation –the Empirical Rule Chebyshev's Theorem		
Measures of Position and Box-and-Whisker Plots – quartiles, interquartile range, percentile, percentile rank, five-number summary, box-and-whisker plots, comparison of data distributions via side-by-side boxplots		
➢ **Probability Theory**		
Experiment, Outcomes, and Sample Space – simple and compound events (just the basic ideas; further treatment is optional)		
Determining Probabilities: Three Conceptual Approaches – definition of the probability of an event, properties of probabilities, classical probability, relative frequency probability, Law of Large Numbers, subjective probability (an informal approach)		
Marginal and Conditional Probabilities – tree diagrams, mutually exclusive events, independent events, complementary events		
Intersection of Events and the Multiplication Rule – contingency tables, joint probability, conditional probability, special cases for independent and mutually exclusive events		**Bayes' Theorem**

Union of Events and the Addition Rule – special case for mutually exclusive events		
		Counting Rules, Factorials and Combinations
➢ **Probability Distributions**		
Random Variables and Probability Distributions of Discrete Random Variables – discrete random variable, continuous random variable, two defining characteristics of a discrete probability distribution, constructing a discrete probability distribution		**Bernoulli distribution, Poisson distribution, geometric and hypergeometric distributions**
Mean and Standard Deviation of a Discrete Random Variable –plus interpretation of the standard deviation (approach can be very informal)		
The Binomial Probability Distribution – the binomial experiment, a trial, a Bernoulli trial, binomial parameters, binomial formula, the shape of the binomial distribution, mean and standard deviation of the binomial distribution		
The Normal Distribution – continuous random variable, continuous probability distribution, shape of a normal probability distribution, calculating probabilities from a normal distribution, assessing normality using probability quantile plots		
The Standard Normal Distribution – z scores, standardizing a normal distribution, calculating probabilities from a standard normal distribution, applications of a normal distribution, determining x- and z-values when an area under a (standard) normal distribution is known.		
		The Normal Approximation to the Binomial – determining the conditions when the binomial distribution can be approximated using a normal distribution, continuity correction factor

➢ **Sampling Distributions of a Statistic**		
Sampling Distributions of a Statistic – population distribution, sampling distribution of the sample mean, sources of variation		
Mean and Standard Deviation of the Sampling Distribution of the Sample Mean		
Shape of the Sampling Distribution of the Sample Mean – from a normally distributed population, from a not normally distributed population, Central Limit Theorem for sample means, applications		
Sampling Distribution of the Sample Proportion – population proportion, sample proportion, mean and standard deviation of the sampling distribution of the sample proportion, shape of the sampling distribution of the sample proportion, Central Limit Theorem for sample proportions, applications		
➢ **Estimation of a Parameter and Tests of Hypotheses for One Population**		
Estimation – point estimate, interval estimate, margin of error, confidence level, confidence interval		
Estimation of a Population Proportion: Large Samples – estimator of the standard deviation of the sampling distribution of a sample proportion, confidence interval for the population proportion		**determining sample size for the estimation of population proportion**
Hypotheses Tests About the Population Proportion: Large Samples – test statistic, p-value **or** critical-value approach		
Estimation of a Population Mean: Population Standard Deviation Known – confidence interval for a population mean, width of a confidence interval,		**determining sample size for estimation of population mean**
Estimation of a Population Mean: Population Standard Deviation Unknown – *t* distribution, degrees of freedom for the t distribution, confidence interval for population mean using the *t* distribution		

Hypotheses Tests – null hypothesis, alternative hypothesis, Type I error, Type II error, significance level of a test, tails of a test, left-tailed test, right-tailed test, two-tailed test		**Choice of:** critical value of a test, rejection region, nonrejection region; **or:** p-value, **power of a test**
		Hypotheses Tests About the Population Mean: Population Standard Deviation Known **–** p-value, p-value approach to hypothesis testing, calculating the p-value of a test, test statistic, calculating the test statistic, critical-value approach to hypothesis testing
Hypotheses Tests About the Population Mean: Population Standard Deviation Unknown **–** test statistic using the t distribution, degrees of freedom, range for the p-value, p-value **or** critical value approach		
➢ **Estimation and Tests of Hypotheses for Two or More Populations**		
Inferences About the Difference Between Two Population Means (Independent Samples) **–** independent samples, dependent samples, sampling distribution of a difference of two sample means, mean and standard deviation of a difference of two sample means, hypothesis test for a difference of two population means, test statistic for a difference of two sample means, p-value approach **or** critical-value approach, confidence interval for a difference of two population means		
CHOICE OF EITHER: **Inferences About the Difference Between Two Population Means (Independent Samples):** σ_1 **and** σ_2 **Unknown but Equal –** pooled standard deviation for two samples, estimator of the standard deviation of a difference of two sample means, hypothesis test for a difference of two population means, degrees of freedom, test statistic using the t distribution for a difference of two sample means, p-value approach **or** critical-value approach		**confidence interval for a difference of population means**

OR: **Inferences About the Difference Between Two Population Means (Independent Samples):** σ_1 **and** σ_2 **Unknown and Unequal –** degrees of freedom, estimator of the standard deviation of a difference of two sample means, hypothesis test for a difference of two population means, test statistic using the t distribution for a difference of two sample means, p-value **or** critical-value approach		**confidence interval for a difference of two population means**
Inferences About the Difference Between Two Population Means (Paired Samples) -paired or matched samples, paired difference, sampling distribution of a difference of two means for paired data, mean and standard deviation of a paired difference, confidence interval for a paired difference, hypothesis testing for a paired difference, test statistic using the t distribution, p-value **or** critical-value approach		
		Inferences About the Difference Between Two Population Proportions for Large and Independent Samples – sampling distribution of a difference of two sample proportions, mean and standard deviation of a difference of two sample proportions, confidence interval for a difference of two population proportions, hypothesis test for a difference of two population proportions, test statistic for a difference of two sample proportions, p-value **or** critical-value approach
		Nonparametric Methods - For example, the Wilcoxon rank sum test and/or the Mann-Whitney U test as a nonparametric alternative to the comparison of the means of two populations based on independent samples and the Wilcoxon signed-rank test for a paired experiment
		Inferences About the Population Variance – confidence interval for a population variance, hypothesis test for a population variance using the Chi-square

		distribution, test statistic for a population variance hypothesis test
Analysis of Variance – the F distribution, one-way analysis of variance testing whether the means of three or more populations are equal, assumptions of the test, degrees of freedom of the test, test statistic using the F distribution, between-samples sums of squares (SSB), within-samples sums of squares (SSW), total sum of squares (SST), variance or mean square between samples (MSB), variance or mean square within samples (MSW), structure of the data (only an informal or intuitive approach – a kind of graphical ANOVA)		
		Goodness-of-Fit Tests – Chi-square distribution, a multinomial experiment, observed frequencies, expected frequencies, degrees of freedom, test statistic using the Chi-square distribution, hypothesis test of goodness-of-fit for categorical variables.
Tests for Independence and Homogeneity – contingency table, test of independence versus test of homogeneity, degrees of freedom, observed and expected frequencies, test statistic for a test of independence and a test of homogeneity using the Chi-square distribution		
➢ **Bivariate Analysis**		
Simple Linear Regression Analysis – explanatory/predictor/independent variables, response/dependent variable, scatter diagram, simple versus multiple regression, regression equation or model, linear regression model versus nonlinear regression model, deterministic model versus probabilistic model, random error term, equation of a regression model, population regression line, observed/actual and estimated/predicted values of the true y-intercept of the population regression line, error sum of squares (SSE), the least squares method, the least squares regression line, positive versus negative linear relationships, assumptions of the regression model, note on regression and causality		**Inferences about the population regression line slope (*B*) –** sampling distribution of the slope (***b***) of the sample regression line, confidence interval of population regression line slope using the t distribution, hypothesis testing of the slope of the population regression line, test statistic using the *t* distribution, *p*-value **or** critical-value approach

		Linear Correlation – linear correlation coefficient , hypothesis testing of the population correlation coefficient, degrees of freedom, test statistic using the sample correlation coefficient (***r***) and the t distribution, critical-value approach, p-value approach
		Standard Deviation of Random Errors and the Coefficient of Determination – degrees of freedom for the regression model, standard deviation of errors, total sum of squares (SST), regression sum of squares (SSR), coefficient of determination
Applications of Correlation and Regression – using the regression model for estimating the mean value of the dependent variable y given an independent value of x on the population regression line		**Using the regression model for predicting a particular value of the dependent variable y on the population regression line, prediction interval**