

Statistics 213

Introduction to Statistics I

(see Course Descriptions for the applicable academic year: <http://www.ucalgary.ca/pubs/calendar/>)

Syllabus

<u>Topics</u>	<u>Number of hours</u>
<p>EXPLORATORY DATA ANALYSIS Populations and sampling, simple random sample. Measures of location and spread: mean, median, mode, variance and standard deviation, quartiles. Grouped data, frequency histograms, shape, symmetry and skewness. The empirical rule. Graphical methods: Stem-and-leaf plots, Box-and-whisker plots.</p>	5
<p>REGRESSION AND CORRELATION Scatter diagrams, linear regression and correlation. Regression predictions - descriptive methods.</p>	2
<p>INTRODUCTION TO PROBABILITY Sample spaces and random events, Venn diagrams. Permutations and combinations. Definitions of probability. Basic probability laws.</p>	4
<p>CONDITIONAL PROBABILITY Conditional probability, probability trees, formula for total probability and Bayes' Theorem. Independence of random events.</p>	3
<p>RANDOM VARIABLES Random variables and their distributions. Joint probability distributions and independence.</p>	3
<p>DISCRETE DISTRIBUTIONS Uniform, binomial, hypergeometric, and Poisson distributions.</p>	3
<p>EXPECTATIONS AND VARIANCES Expectations, variances and covariances of random variables and their linear combinations. Application to discrete random variables.</p>	6
<p>CONTINUOUS RANDOM VARIABLES Concept of continuous random variable and its distribution. Uniform and Normal distributions and their properties. Central Limit Theorem. Approximation of the binomial by the Normal distribution. <u>Optional</u>: exponential distribution.</p>	5
<p>SAMPLING DISTRIBUTIONS Distribution of the sample mean from a Normal population. Central Limit Theorem and large sample mean distribution. Distribution of the sample proportion.</p>	2
<p>ESTIMATION AND HYPOTHESIS TESTING Introduction to confidence intervals and hypothesis testing using simple examples involving means and proportions.</p>	3
TOTAL HOURS	36

Course Outcomes

Students successfully completing this course should be able to:

1. Identify the population of interest, or the target population. Differentiate between the population and the sample; differentiate between a parameter and a statistic.
2. Make the distinction between a quantitative and qualitative variable. Explain the three different properties of any population variable: the distribution shape, the center of the distribution, and the spread of the distribution. Construct various graphical techniques to make conclusions of the shape of the underlying distribution, the different measures of center and dispersion. Compare the concepts of percentiles and quartiles, and what they mean with regards to the population of interest.
3. Compute the probabilities of simple and compound events. Give examples of the concepts of mutually exclusive events, independent events, and conditional events. Illustrate how an event can be transformed into a real number through the use of random variables; show that a random variable has a distribution, with a measure of center and a measure of dispersion.
4. Compute the expected value, the variance and the standard deviation of a generic discrete and continuous random variable. Compute the expected total and its standard deviation of a linear function of certain random variables.
5. Illustrate that certain random events can be described by probability models. Differentiate between the probability models (the Binomial, Poisson, Uniform/Exponential, Hypergeometric and Normal distributions) and apply each to find probabilities. Find a percentile under the Normal distribution. A knowledge of each distribution-shape, measure of center, and measure of dispersion-is also expected.
6. Describe the Central Limit and apply to both the sample mean and sample proportion to determine how likely they are to fall within a given range of values.
7. Take a bivariate data set and (i) determine the strength of a linear relationship between the two variables of interest based on a scatter plot and the correlation coefficient, (ii) build a simple linear regression line and interpret the meaning of the slope and intercept parameter estimates, (iii) outline and check assumptions behind the simple linear model, and (iv) find the coefficient of determination and explain its meaning.
8. Construct and interpret the confidence interval for a population mean and a population proportion. Confidence interval estimation of the population mean will emphasize the use of the Student's T-distribution.
9. Compute the required sample size for a given confidence level and tolerable amount of sampling error when the statistical investigation involves estimation of either a population mean or a population proportion.
10. Execute statistical hypothesis testing for a population mean and a population proportion. This includes (i) set up the statistical null and alternative hypotheses (ii) identify the appropriate version of the test statistic and compute the value of this test statistic, (iii) state the rejection region, calculate the P-value, (iv) tell whether the data supports the null hypothesis or not, and (vi) interpret the meaning of the P-value in the context of the data. That is, describe the event that the P-value finds the probability of.

* * * * *