Mathematics of Data Management, Grade 12, University Preparation (MDM4U)

Organization of Data for Analysis

Organizing Data

OD1.01 locate data to answer questions of significance or personal interest, by searching well-organized databases;
OD1.02 use the Internet effectively as a source for databases;
OD1.03 create database or spreadsheet templates that facilitate the manipulation and retrieval of data from large bodies of information that have a variety of characteristics (e.g., a compact disc collection classified by artist, by date, by type of music).

Using Diagrams to Solve Problems

OD2.01 represent simple iterative processes (e.g., the water cycle, a person's daily routine, the creation of a fractal design), using diagrams that involve branches and loops;
OD2.02 represent complex tasks (e.g., searching a list by using algorithms; classifying organisms; calculating dependent or independent outcomes in probability) or issues (e.g., the origin of global warming), using diagrams (e.g., tree diagrams, network diagrams, cause-and-effect diagrams, time lines);
OD2.03 solve network problems (e.g., scheduling problems, optimum-path problems, critical-path problems), using introductory graph theory.

Using Matrices to Model and Solve Problems

OD3.01 represent numerical data, using matrices, and demonstrate an understanding of terminology and notation related to matrices;
OD3.02 demonstrate proficiency in matrix operations, including addition, scalar multiplication, matrix multiplication, the calculation of row sums, and the calculation of column sums, as necessary to solve problems, with and without the aid of technology;
OD3.03 solve problems drawn from a variety of applications (e.g., inventory control, production costs, codes), using matrix methods

Counting and Probability

Solving Counting Problems

CP1.01 use Venn diagrams as a tool for organizing information in counting problems;
CP1.02 solve introductory counting problems involving the additive and multiplicative counting principles;
CP1.03 express the answers to permutation and combination problems, using standard combinatorial symbols, [e.g., , P(n, r)];
CP1.04 evaluate expressions involving factorial notation, using appropriate methods (e.g., evaluating mentally, by hand, by using a calculator);
CP1.05 solve problems, using techniques for counting permutations where some objects may be alike;
CP1.06 solve problems, using techniques for counting combinations;
CP1.07 identify patterns in Pascal's triangle and relate the terms of Pascal's triangle to values of , to the expansion of a binomial, and to the solution of related problems (*Sample problem:* A girl's school is 5 blocks west and 3 blocks south of her home. Assuming that she leaves home and walks only west or south, how many different routes can she take to school?);
CP1.08 communicate clearly, coherently, and precisely the solutions to counting problems.

Determining and Interpreting Theoretical Probabilities

CP2.01 solve probability problems involving combinations of simple events, using counting techniques [i.e., $P(A \text{ or } B)$, $P(A \text{ and } B)$, and $P(A)$];

CP2.02 identify examples of discrete random variables (e.g., the sums that are possible when two dice are rolled);
CP2.03 construct a discrete probability distribution function by calculating the probabilities of a discrete random variable;
CP2.04 calculate expected values and interpret them within applications (e.g., lottery prizes, tests of the fairness of games, estimates of wildlife populations) as averages over a large number of trials;
CP2.05 determine probabilities, using the binomial distribution (*Sample problem:* A light-bulb manufacturer estimates that 0.5% of the bulbs manufactured are defective. If a client places an order for 100 bulbs, what is the probability that at least one bulb is defective?);
CP2.06 interpret probability statements, including statements about odds, from a variety of sources.

Simulating and Predicting

CP3.01 identify the advantages of using simulations in contexts;

CP3.02 design and carry out simulations to estimate probabilities in situations for which the calculation of the theoretical probabilities is difficult or impossible (*Sample problem:* A set of 6 baseball cards can be collected from cereal boxes. If the different cards are evenly distributed throughout the boxes, carry out a simulation to determine the probability of collecting one complete set in a purchase of 14 boxes);

CP3.03 assess the validity of some simulation results by comparing them with the theoretical probabilities, using the probability concepts developed in the course (*Sample problem:* A light-bulb manufacturer estimates that 0.5% of the bulbs manufactured are defective. Carry out a simulation to estimate the probability that at least one bulb is defective in an order of 100 bulbs).

Statistics

Collecting Data

ST1.01 demonstrate an understanding of the purpose and the use of a variety of sampling techniques (e.g., a simple random sample, a systematic sample, a stratified sample);

ST1.02 describe different types of bias that may arise in surveys (e.g., response bias, measurement bias, non-response bias, sampling bias);

ST1.03 illustrate sampling bias and variability by comparing the characteristics of a known population with the characteristics of samples taken repeatedly from that population, using different sampling techniques;

ST1.04 organize and summarize data from secondary sources (e.g., the Internet, computer databases), using technology (e.g., spreadsheets, graphing calculators).

Analysing Data Involving One Variable

ST2.01 compute, using technology, measures of one-variable statistics (i.e., the mean, median, mode, range, interquartile range, variance, and standard deviation), and demonstrate an understanding of the appropriate use of each measure;

ST2.02 interpret one-variable statistics to describe characteristics of a data set;

ST2.03 describe the position of individual observations within a data set, using *z*-scores and percentiles.

Solving Problems Involving the Normal Distribution

ST3.01 identify situations that give rise to common distributions (e.g., bimodal, U-shaped, exponential, skewed, normal);

ST3.02 demonstrate an understanding of the properties of the normal distribution (e.g., the mean, median, and mode are equal; the curve is symmetric about the mean; 68% of the population are within one standard deviation of the mean) and use these properties to solve problems;

ST3.03 make probability statements about normal distributions, on the basis of information drawn from a variety of applications.

Describing the Relationship Between Two Variables

ST4.01 define the correlation coefficient as a measure of the fit of a scatter graph to a linear model;

ST4.02 calculate the correlation coefficient for a set of data, using graphing calculators or statistical software;

ST4.03 demonstrate an understanding of the distinction between cause-effect relationships and the mathematical correlation between variables;

ST4.04 describe possible misuses of regression (e.g., use with too small a sample, use without considering the effect of outliers, inappropriate extrapolation).

Evaluating Validity

ST5.01 explain examples of the use and misuse of statistics in the media;

ST5.02 assess the validity of conclusions made on the basis of statistical studies, by analysing possible sources of bias in the studies (e.g., sampling bias) and by calculating and interpreting additional statistics, where possible (e.g., measures of central tendency, the standard deviation);

ST5.03 explain the meaning and the use in the media of indices based on surveys (e.g., the consumer price index, the cost of living index).

Integration of the Techniques of Data Management

Carrying Out a Culminating Project

DM1.01 pose a significant problem whose solution would require the organization and analysis of a large amount of data;
DM1.02 select and apply the tools of the course (e.g., methods for organizing data, methods for calculating and interpreting measures of probability and statistics, methods for data collection) to design and carry out a study of the problem;
DM1.03 compile a clear, well-organized, and fully justified report of the investigation and its findings.

Presenting and Critiquing Projects

DM2.01 create a summary of a project to present within a restricted length of time, using communications technology effectively;
DM2.02 answer questions about a project, fully justifying mathematical reasoning;
DM2.03 critique the mathematical work of others in a constructive fashion.