

"For You light my lamp; the LORD my God illumines my darkness.
For by You I can run upon a troop; and by my God I can leap over a wall.
As for God, His way is blameless; the word of the LORD is tried;
He is a shield to all who take refuge in Him."
-- Psalm 18:28-30

- Please show all your work! No partial credit will be given for incorrect answers with no work shown. Please draw a box around your final answer.
- Calculators are permitted, but no notes, text, laptops, PDAs, or electronic dictionaries. Cell phones should be muted and left in your pocket or bag.
- Table 3 is attached to the back. You may detach it for your reference.

1. Classify each of the following statements (which may or may not be true) as either **(D)escriptive** or **(I)nfere[n]tial**: **[4]**

- (a) 19.2% of BC residents in 2004 were clinically obese.
- (b) Of the children and adolescents studied in the 2004 Canadian Community Health Survey, 8% were clinically obese.
- (c) About 90% of students in our MATH108 class are nursing students.
- (d) 30.1% of male registered nurses (RNs) are under the age of 40.

2. Mark each of the following **variables** as nominal (N), ordinal (O), discrete (D), or continuous (C): **[5]**

- (a) How many children are in a family
- (b) Satisfaction with current family doctor, rated as "Very Dissatisfied", "Dissatisfied", "Satisfied", or "Very Satisfied"
- (c) Whether a student passes a course or not
- (d) Hemoglobin count, in g/dL
- (e) Age, divided into categories "0-10", "11-18", "19-25", "26-39", "40-59", "60 and up"

3. The home provinces of 16 students in a class are listed below. Draw a **Pareto** chart showing the distribution of home province in this sample. **[4]**

AB, BC, BC, SK, AB, SK, BC, ON, SK, BC, BC, AB, BC, ON, MB, SK

4. The table below lists the number of purple starthistle plants (a noxious invasive weed in WA state) in nine equally-sized plots.
- 41, 57, 47, 32, 42, 28, 36, 53, 42
- (a) Construct a relative frequency **histogram**, classifying the data by bins of width 10 plants/plot. **[4]**
- (b) Find the sample **mean**. Show your work. **[2]**
- (c) Find the **mode** and the **midrange**. **[2]**
- (d) Find the sample **standard deviation**. Show your work. **[4]**
- (e) Draw a **boxplot** for the data. Show your work. **[4]**
5. Say that in a certain town 63% of the people vote Conservative, and the probability that a townspeople drives a large truck is $\frac{3}{7}$. If you pick a random townspeople, there is a 27% chance that the townspeople is a Conservative driving a large truck.
- (a) What fraction of Conservatives in this town drive large trucks? **[3]**
- (b) In this town, is voting Conservative **mutually exclusive** of driving a large truck? Why or why not? **[1]**
- (c) In this town, is voting Conservative **independent** of driving a large truck? Why or why not? Interpret what this means in the context of the townspeople. **[3]**

(d) The town council is a randomly-selected group of 6 townspeople. What is the chance that exactly **half** of the town council votes Conservative? [3]

(e) What is the chance that a **majority** of the town council votes Conservative? [4]

(f) The small business association is a randomly-selected group of 50 townspeople. What is the chance that more than **half** of the small business association votes Conservative? [4]

6. An assay (test) measuring inorganic phosphate in blood is imprecise, with values normally distributed around the true phosphate concentration, with a standard deviation of 0.2 mmol/L. Blood phosphate values of over 1.8 mmol/L are considered "high".

(a) What is the probability that this assay returns a result **within** ± 0.25 mmol/L of the true value? [2]

(b) If the true concentration of phosphate is 1.63 mmol/L (not "high"), what is the probability that the assay could still return a value that is considered "**high**"? [2]

(c) The standard deviation is just one way of measuring the dispersion (imprecision) of the assay results. The **interquartile range (IQR)** is another. Find the IQR for this assay. [3]

7. A particular screening test for breast cancer has a 11% false-positive rate (i.e., 89% **specificity**) and a 7% false-negative rate (i.e., 93% **sensitivity**).
- (a) Suppose the test is applied to a group of patients, 40% of whom are known to have breast cancer. Draw an **event tree** for the outcomes of the test. Label the tree with probabilities for each branch. Also calculate the probabilities of each final outcome (leaf of the tree). [4]
- (b) What is the probability that a random patient from this group will **test positive** for breast cancer using this screening test? [2]
- (c) What is the probability that a patient from this group who tested positive **actually** has the disease? [2]
8. Suppose that when a job candidate comes to interview for a staff position at TWU, the probability that he or she will want the job (A) after the interview is 72%. The probability that TWU wants the candidate (B) is 40%. Also, $P(A|B)=0.80$.
- (a) Find $P(A \cap B)$, draw a **Venn** diagram, and **interpret** it in words. [3]
- (b) Find $P(B | A)$, and **interpret** it in words. [3]
- (c) Are events A and B **independent**? Explain. Interpret (in the context of job interviews) what it means for A and B to be independent. [2]