Key Formulas for STA1 Q1 & Q2

Chapter 2 - Organizing and Graphing Data

- Relative frequency of a class $= f/\Sigma f$
- Percentage of a class = (Relative frequency) $\times$ 100
- Class midpoint or mark = (Upper limit+Lower limit)/2
- Class width = Upper boundary - Lower boundary
- Cumulative relative frequency

$$= \frac{\text{Cumulative frequency}}{\text{Total observations in the data set}}$$

- Cumulative percentage

$$= (\text{Cumulative relative frequency}) \times 100$$

Chapter 3 - Numerical Descriptive Measures

- Mean for ungrouped data: $\mu = \Sigma x/N$ and $\bar{x} = \Sigma x/n$
- Mean for grouped data: $\mu = \Sigma mf/N$ and $\bar{x} = \Sigma mf/n$, where m is the midpoint and f is the frequency of a class
- Median for ungrouped data
  = Value of the middle term in a ranked data set
- Range = Largest value - Smallest value
- Variance for ungrouped data:

$$\sigma^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}{N} \quad \text{and} \quad s^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}$$

  where σ^2 is the population variance and s^2 is the sample variance

- Standard deviation for ungrouped data:

$$\sigma = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}{N}} \quad \text{and} \quad s = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1}}$$

  where σ and s are the population sample standard deviations, respectively

- Variance for grouped data:

$$\sigma^2 = \frac{\Sigma m^2 f - \frac{(\Sigma mf)^2}{N}}{N} \quad \text{and} \quad s^2 = \frac{\Sigma m^2 f - \frac{(\Sigma mf)^2}{n}}{n-1}$$

- Standard deviation for grouped data:

$$\sigma = \sqrt{\frac{\Sigma m^2 f - \frac{(\Sigma mf)^2}{N}}{N}} \text{ and } s = \sqrt{\frac{\Sigma m^2 f - \frac{(\Sigma mf)^2}{n}}{n-1}}$$

Chapter 4 - Probability

- Classical probability rule for a simple event:

$$P(E_i) = \frac{1}{\text{Total number of outcomes}}$$

- Classical probability rule for a compound event:

$$P(A) = \frac{\text{Number of outcomes in A}}{\text{Total number of outcomes}}$$

- Relative frequency as an approximation of probability:

$$P(A) = \frac{f}{n}$$

- Conditional probability of an event:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \text{ and } P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

- Condition for independence of events:

$$P(A) = P(A|B) \quad \text{and/or} \quad P(B) = P(B|A)$$

- For complementary events: $P(A) + P(\bar{A}) = 1$
- Multiplication rule for dependent events:

$$P(A \text{ and } B) = P(A)P(B|A)$$

- Multiplication rule for independent events:

$$P(A \text{ and } B) = P(A)P(B)$$

- Joint probability of two mutually exclusive events:

$$P(A \text{ and } B) = 0$$

- Addition rule for mutually nonexclusive events:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

- Addition rule for mutually exclusive events:

$$P(A \text{ or } B) = P(A) + P(B)$$

Chapter 5 - Discrete Random Variables and Their Probability Distribution

- Mean of a discrete random variable x: $\mu = \sum xP(x)$
- Standard deviation of a discrete random variable x:

$$\sigma = \sqrt{(\sum x^2 P(x)) - \mu^2}$$

Chapter 6 - Continuous Random Variables and the Normal Distribution

- z value for an x value: $z = \frac{x-\mu}{\sigma}$
- Value of x when $\mu, \sigma,$ and z are known: $x = \mu + z\sigma$

Chapter 13 - Simple Linear Regression

- Simple linear regression model: $y = A + Bx + \in$
- Estimated simple linear regression model: $\hat{y} = a + bx$

Chapter 14 - Multiple Regression

- Multiple regression model:
  $y = A + B_1x_1 + B_2x_2 + B_3x_3 + \cdots + B_kx_k + \in$
- Estimated multiple regression model:
  $\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_kx_k$