Midterm-Friendly Formula Sheet

Descriptive Statistics

Find Unknown Position from Known Percentile: To find the the position of the data point that is the kth percentile in a data set, use the formula

$$i = \frac{k}{100}(n+1)$$

Then round the position i both down and up to the nearest integers. The average of the values at these positions is your kth percentile.

Find Unknown Percentile from Known Position: To find the percentile of a data point from its position:

1) Order the data from smallest to largest.
2) Let x = the number of data values counting from the bottom of the data list up to but not including the data value for which you want to find the percentile.
3) Let y = the number of data values equal to the data value for which you want to find the percentile.
4) Let n = the total number of data points.
5) Then the percentile value of the data point is found by computing

$$p = \left(\frac{x + 0.5y}{n}\right) \cdot 100$$

and rounding to the nearest integer.

Sample Mean: $\overline{x} = \dfrac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \dfrac{\sum_{j=1}^{n} x_j}{n}$. (Add up all data values and divide by the total number of data values.)

Median: Line up data values from smallest to largest. The median is the middle value. If the data set has an even number of values, the median is the average of the 2 middle values.

Mode: The mode is the data value which occurs most often.

Quartile 1: The median of the lower-half of the data set (all the data values below the median).

Quartile 3: The median of the upper-half of the data set (all the data values above the median).

Five Number Summary: Minimum, Quartile 1, Median, Quartile 3, Maximum

Sample Standard Deviation: $s = \sqrt{\dfrac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2}{n-1}} = \sqrt{\dfrac{\sum_{j=1}^{n}(x_j - \overline{x})^2}{n-1}}$

z-Scores: $z = \dfrac{x - \mu}{\sigma}$ where μ is a population mean and σ is a population standard deviation.

Basic Probability

Conditional Probability: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

General OR Event: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

General AND Event: $P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

Addition Rule for Mutually Exclusive Events: $P(A \cup B) = P(A) + P(B)$

Multiplication Rule for Independent Events: $P(A \cap B) = P(A) \cdot P(B)$

General Random Variables

Mean: $\mu = \displaystyle\sum_{j=1}^{n} x_j P(X = x_j) = x_1 P(X = x_1) + x_2 P(X = x_2) + \cdots + x_n P(X = x_n)$

Standard Deviation: $\sigma = \sqrt{\displaystyle\sum_{j=1}^{n} (x_j - \mu)^2 P(X = x_j)}$
$= \sqrt{(x_1 - \mu)^2 P(X = x_1) + (x_2 - \mu)^2 P(X = x_2) + \cdots + (x_n - \mu)^2 P(X = x_n)}$

Binomial Distribution

PDF	**CDF**	**Mean μ**	**Standard Deviation σ**
$P(X = k) = \binom{n}{k} p^k q^{n-k}$	Use Technology	$\mu = np$	$\sigma = \sqrt{npq}$