STAT 305 Quiz II
Reference Sheet

Numeric Summaries

mean	$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$
population variance	$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$
population standard deviation	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$
sample variance	$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$
sample standard deviation	$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$

Functions

Quantile Function $Q(p)$ For a univariate sample consisting of n values that are ordered so that $x_1 \leq x_2 \leq \ldots \leq x_n$ and value p where $0 \leq p \leq 1$, let $i = \lfloor n \cdot p + 0.5 \rfloor$. Then the quantile function at p is:

$$Q(p) = \begin{cases} x_i & \lfloor n \cdot p + .5 \rfloor = n \cdot p + .5 \\ x_i + (np - i + .5)(x_{i+1} - x_i) & \lfloor n \cdot p + .5 \rfloor \neq n \cdot p + .5 \end{cases}$$

Measures of Central Tendency

- $Q\left(\frac{1-.5}{n}\right)$ is called the **minimum** and $Q\left(\frac{n-.5}{n}\right)$ is called the **maximum** of a distribution.

- $Q(.5)$ is called the **median** of a distribution.

- $Q(.25)$ and $Q(.75)$ are called the **first (or lower) quartile** and **third (or upper) quartile** of a distribution, respectively.

- The **interquartile range (IQR)** is defined as

$$IQR = Q(.75) - Q(.25)$$

.

- An **outlier** is a data point that is larger than $Q(.75) + 1.5 * IQR$ or smaller than $Q(.25) - 1.5 * IQR$.