Notation, Terms, and Equations for Statistical Methods for Managers

Ken Kelley's MSBA Notes

The following helps to summarize some of the notation, terms, and equations as used in the notes and discussed in class. Though, this is not a complete list and some terms may be used differently depending on context.

Terms and Notation

- Frequency is the number (i.e., the count) of occurrences (e.g., in a particular category or in a particular group).
- Relative frequency (e.g., in a category) is the particular frequency divided by the total frequency. The relative frequency gives the proportion of the total number in a particular group or category.
- x_i, y_i, z_i are the variables x, y, and z, respectively, representing the realized value for a particular entity on the variable.
- n is the sample size of a sample ($i = 1, \ldots, n$).
- N is the population size (which may not be known).
- $\sum_{i=1}^{n}$ is the summation operator, which denotes summing what is to the right of the summation from the ith to the nth observation (e.g., $\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$).
- $\bar{x}$ is the sample mean.
- s^2 is the estimated variance and s the estimated standard deviation.
- $s_{\bar{x}}$ is the estimated standard deviation of the mean (i.e., the standard error of the mean).
- z_i is the value of the z-score for the ith individual.
- z is the value of a test statistic.
- t is the value of a test statistic.
- $z_{1-\alpha/2}$ is the critical value of a z test statistic at the $1 - \alpha/2$ quantile (e.g, .975).
- $t_{1-\alpha/2}$ is the critical value of a t test statistic at the $1 - \alpha/2$ quantile (e.g, .975).
- $F_{1-\alpha}$ is the critical value of an F test statistic at the $1 - \alpha$ quantile (e.g, .95).
- r_{xy} is the correlation coefficient of variables x and y (ρ_{xy} is the population value).
- s_{xy} is the covariance of variables x and y (σ_{xy} is the population value).
- μ_0 is the value of the mean under the specified null hypothesis that is to be tested.

Equations

- $\bar{x} = \frac{\sum x_i}{n} = \frac{x_1+x_2+\cdots+x_n}{n}$.
- For grouped data, $\bar{x} = \frac{\sum f_i x_i}{n}$.
- Median = The score in the $\frac{n+1}{2}$ position, if scores are arranged from lowest to highest.
- 1st Quartile = The 25th percentile.
- 3rd Quartile = The 75th percentile.
- Interquartile Range = 3rd Quartile - 1st Quartile.
- $s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$.
- $s = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}} = \sqrt{s^2}$.
- $z_i = \frac{x_i-\mu}{\sigma}$.
- From a z-score to the scale of x: $x_i = \sigma * z_i + \mu$.
- $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.
- $s_{\bar{x}} = \frac{s}{\sqrt{n}}$.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$.
- $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$.
- MOE (for z)=$\sigma_{\bar{x}} \times z_{Critical\ Value}$.
- MOE (for t)=$s_{\bar{x}} \times t_{Critical\ Value}$.
- $s_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{n-1}$.
- $r = \frac{s_{xy}}{s_x s_y}$.
- $P(A) = \frac{A}{A+B}$, where A and B are different events.
- If events A and B are mutually exclusive: $P(A \cup B) = P(A \text{ or } B) = P(A) + P(B)$.
- If events A and B are not mutually exclusive: $P(A \cup B) = P(A) + P(B) - P(A \text{ and } B)$.
- If events A and B are independent, then $P(A \cap B) = P(A) \times P(B)$. Additionally, if A and B are independent, then $P(A|B) = P(A)$.
- Conditional Probabilities of A given the occurrence of B: $P(A|B) = \frac{P(A \cap B)}{P(B)}$.
- Odds of A = $\frac{P(A)}{1-P(A)}$.