Statistics Formulas

Statistics is essentially the process of learning from data. The goal of statistics is to make correct statements or inferences about a population based on a sample.

This handout consists of terminology and formulas contained in *Essentials of Statistics* (4th edition) by Mario Triola.

Chapter 3 – *Descriptive Statistics*:

$\bar{x} = \frac{\sum x}{n}$ *Mean of a sample*

$\mu = \frac{\sum x}{N}$ *Mean of a population*

$\bar{x} = \frac{\sum (x \cdot f)}{\sum f}$ *Mean of a frequency distribution*

$s = \sqrt{\frac{\sum (x-\bar{x})^2}{(n-1)}}$ *Standard deviation of a sample*

$s = \sqrt{\frac{n\left(\sum x^2\right) - \left(\sum x\right)^2}{n(n-1)}}$ *Standard deviation (shortcut)*

$s = \sqrt{\frac{n\left[\sum \left(f \bullet x^2\right)\right] - \left[\sum \left(f \bullet x\right)\right]^2}{n(n-1)}}$ *Standard deviation (frequency)*

s^2 *(Sample) variance*

σ^2 *(Population) variance*

Chapter 4 – *Probabilities*:

$P(A \text{ or } B) = P(A) + P(B)$
if A, B are mutually exclusive

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
if A and B are not mutually exclusive

$P(A \text{ and } B) = P(A) \bullet P(B)$ *if A, B are independent*

$P(A \text{ and } B) = P(A) \bullet P(B|A)$ *if A, B are dependent*

${}_nP_r = \frac{n!}{(n-r)!}$ *Permutations (no elements alike)*

${}_nP_r = \frac{n!}{n_1! n_2! \ldots n_k!}$ *Permutations (n_1 alike, …)*

${}_nC_r = \frac{n!}{(n-r)! r!}$ *Combinations*

Chapter 5 – *Probability Distributions*:

$\mu = \sum \left[x \bullet P(x)\right]$ *Mean probability distribution*

$\sigma = \sqrt{\sum \left[x^2 \bullet P(x)\right] - \mu^2}$ *Standard deviation (probability dist)*

$P(x) = \frac{n!}{(n-x)! x!} \bullet p^x q^{n-x}$ *Binomial probability*

$\mu = n \bullet p$ *Mean (binomial)*

$\sigma^2 = n \bullet p \bullet q$ *Variance (binomial)*

$\sigma = \sqrt{n \bullet p \bullet q}$ *Standard deviation (binomial)*

Statistics Formulas

Chapter 6 –*Normal Distribution*:

$z = \frac{x - \bar{x}}{s}$ or $\frac{x - \mu}{\sigma}$ *standard score*

$\mu_{\bar{x}} = \mu$ *central limit theorem*

$x = zs + \bar{x}$

or *standard score* *(different form)*

$x = z\sigma + \mu$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ *central limit theorem (standard error)*

Chapter 7 –*Confidence Intervals*:

$\hat{p} - E < p < \hat{p} + E$ *proportion* *where* $E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

$\bar{x} - E < \mu < \bar{x} + E$ *mean* *where* $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ *if* σ *is known or* $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$ *if* σ *is unknown*

Chapter 7 –*Sample Size Determination*:

$n = \frac{[z_{\alpha/2}]^2 \bullet 0.25}{E^2}$ *proportion (* $\hat{p}$ *and* $\hat{q}$ *are not known)*

$n = \frac{[z_{\alpha/2}]^2\, \hat{p}\hat{q}}{E^2}$ *proportion (* $\hat{p}$ *and* $\hat{q}$ *are known)*

$n = \left[\frac{z_{\alpha/2}\sigma}{E}\right]^2$ *mean*

Chapter 8 – *Test Statistics (one population)*:

$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$ *proportion – one population*

$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ *mean – one population*

$(\sigma \quad unknown)$

$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ *mean – one population*

$(\sigma \quad known)$