Formula/Table Card for Weiss's *Introductory Statistics, 9/e*

Larry R. Griffey

Notation

n = sample size
$\bar{x}$ = sample mean
s = sample stdev
Q_j = jth quartile
N = population size
μ = population mean
σ = population stdev
d = paired difference
$\hat{p}$ = sample proportion
p = population proportion
O = observed frequency
E = expected frequency

Chapter 3 Descriptive Measures

- Sample mean: $\bar{x} = \frac{\Sigma x_i}{n}$
- Range: Range = Max − Min
- Sample standard deviation: Also $s^2 = \frac{\Sigma f x^2 - (\Sigma f x)^2/n}{n-1}$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} \quad \text{or} \quad s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$$

- Interquartile range: $\text{IQR} = Q_3 - Q_1$
- Lower limit = $Q_1 - 1.5 \cdot \text{IQR}$, Upper limit = $Q_3 + 1.5 \cdot \text{IQR}$
- Population mean (mean of a variable): $\mu = \frac{\Sigma x_i}{N}$
- Population standard deviation (standard deviation of a variable):

$$\sigma = \sqrt{\frac{\Sigma(x_i - \mu)^2}{N}} \quad \text{or} \quad \sigma = \sqrt{\frac{\Sigma x_i^2}{N} - \mu^2}$$

- Standardized variable: $z = \frac{x - \mu}{\sigma}$

Chapter 4 Probability Concepts

- Probability for equally likely outcomes:

$$P(E) = \frac{f}{N}$$

where f denotes the number of ways event E can occur and N denotes the total number of outcomes possible.

- Special addition rule:

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

($A, B, C, \ldots$ mutually exclusive)

- Complementation rule: $P(E) = 1 - P(\text{not } E)$
- General addition rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \,\&\, B)$
- Conditional probability rule: $P(B \mid A) = \frac{P(A \,\&\, B)}{P(A)}$
- General multiplication rule: $P(A \,\&\, B) = P(A) \cdot P(B \mid A)$
- Special multiplication rule:

$$P(A \,\&\, B \,\&\, C \,\&\, \cdots) = P(A) \cdot P(B) \cdot P(C) \cdots$$

($A, B, C, \ldots$ independent)

- Rule of total probability:

$$P(B) = \sum_{j=1}^{k} P(A_j) \cdot P(B \mid A_j)$$

($A_1, A_2, \ldots, A_k$ mutually exclusive and exhaustive)

- Bayes's rule:

$$P(A_i \mid B) = \frac{P(A_i) \cdot P(B \mid A_i)}{\sum_{j=1}^{k} P(A_j) \cdot P(B \mid A_j)}$$

($A_1, A_2, \ldots, A_k$ mutually exclusive and exhaustive)

- Factorial: $k! = k(k-1)\cdots 2 \cdot 1$
- Permutations rule: ${}_mP_r = \frac{m!}{(m-r)!}$
- Special permutations rule: ${}_mP_m = m!$
- Combinations rule: ${}_mC_r = \frac{m!}{r!(m-r)!}$
- Number of possible samples: ${}_NC_n = \frac{N!}{n!(N-n)!}$

Chapter 5 Discrete Random Variables

- Mean of a discrete random variable X: $\mu = \Sigma x P(X = x)$
- Standard deviation of a discrete random variable X:

$$\sigma = \sqrt{\Sigma(x-\mu)^2 P(X = x)} \quad \text{or} \quad \sigma = \sqrt{\Sigma x^2 P(X = x) - \mu^2}$$

- Factorial: $k! = k(k-1)\cdots 2 \cdot 1$
- Binomial coefficient: $\binom{n}{x} = \frac{n!}{x!(n-x)!}$
- Binomial probability formula:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where n denotes the number of trials and p denotes the success probability.

- Mean of a binomial random variable: $\mu = np$
- Standard deviation of a binomial random variable:

$$\sigma = \sqrt{np(1-p)}$$

- Poisson probability formula: $P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$
- Mean of a Poisson random variable: $\mu = \lambda$
- Standard deviation of a Poisson random variable: $\sigma = \sqrt{\lambda}$

Chapter 6 The Normal Distribution

- z-score for an x-value: $z = \frac{x - \mu}{\sigma}$
- x-value for a z-score: $x = \mu + z \cdot \sigma$

Formula/Table Card for Weiss's *Introductory Statistics*, 9/e

Larry R. Griffey

Chapter 7 The Sampling Distribution of the Sample Mean

- Mean of the variable $\bar{x}$: $\mu_{\bar{x}} = \mu$
- Standard deviation of the variable $\bar{x}$: $\sigma_{\bar{x}} = \sigma/\sqrt{n}$

Chapter 8 Confidence Intervals for One Population Mean

- Standardized version of the variable $\bar{x}$:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- z-interval for μ (σ known, normal population or large sample):

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

- Margin of error for the estimate of μ: $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
- Sample size for estimating μ:

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$$

rounded up to the nearest whole number.

- Studentized version of the variable $\bar{x}$:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- t-interval for μ (σ unknown, normal population or large sample):

$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

with df $= n - 1$.

Chapter 9 Hypothesis Tests for One Population Mean

- z-test statistic for H_0: $\mu = \mu_0$ (σ known, normal population or large sample):

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

- t-test statistic for H_0: $\mu = \mu_0$ (σ unknown, normal population or large sample):

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

with df $= n - 1$.

- Symmetry property of a Wilcoxon signed-rank distribution:

$$W_{1-A} = n(n + 1)/2 - W_A$$

- Wilcoxon signed-rank test statistic for H_0: $\mu = \mu_0$ (symmetric population):

$$W = \text{sum of the positive ranks}$$

Chapter 10 Inferences for Two Population Means

- Pooled sample standard deviation:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- Pooled t-test statistic for H_0: $\mu_1 = \mu_2$ (independent samples, normal populations or large samples, and equal population standard deviations):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{(1/n_1) + (1/n_2)}}$$

with df $= n_1 + n_2 - 2$.

- Pooled t-interval for $\mu_1 - \mu_2$ (independent samples, normal populations or large samples, and equal population standard deviations):

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \cdot s_p\sqrt{(1/n_1) + (1/n_2)}$$

with df $= n_1 + n_2 - 2$.

- Degrees of freedom for nonpooled t-procedures:

$$\Delta = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

rounded down to the nearest integer.

- Nonpooled t-test statistic for H_0: $\mu_1 = \mu_2$ (independent samples, and normal populations or large samples):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$$

with df $= \Delta$.

- Nonpooled t-interval for $\mu_1 - \mu_2$ (independent samples, and normal populations or large samples):

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \cdot \sqrt{(s_1^2/n_1) + (s_2^2/n_2)}$$

with df $= \Delta$.

- Symmetry property of a Mann–Whitney distribution:

$$M_{1-A} = n_1(n_1 + n_2 + 1) - M_A$$

- Mann–Whitney test statistic for H_0: $\mu_1 = \mu_2$ (independent samples and same-shape populations):

$$M = \text{sum of the ranks for sample data from Population 1}$$

- Paired t-test statistic for H_0: $\mu_1 = \mu_2$ (paired sample, and normal differences or large sample):

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

with df $= n - 1$.

Formula/Table Card for Weiss's *Introductory Statistics, 9/e*

Larry R. Griffey

- Paired t-interval for $\mu_1 - \mu_2$ (paired sample, and normal differences or large sample):

$$\bar{d} \pm t_{\alpha/2} \cdot \frac{s_d}{\sqrt{n}}$$

with df = $n - 1$.

Paired Wilcoxon signed-rank test statistic for H_0: $\mu_1 = \mu_2$ (paired sample and symmetric differences):

$$W = \text{sum of the positive ranks}$$

Chapter 14 Descriptive Methods in Regression and Correlation

- S_{xx}, S_{xy}, and S_{yy}:

$$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - (\Sigma x_i)^2/n$$

$$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)/n$$

$$S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - (\Sigma y_i)^2/n$$

- Regression equation: $\hat{y} = b_0 + b_1 x$, where

$$b_1 = \frac{S_{xy}}{S_{xx}} \quad \text{and} \quad b_0 = \frac{1}{n}(\Sigma y_i - b_1 \Sigma x_i) = \bar{y} - b_1\bar{x}$$

- Total sum of squares: $SST = \Sigma(y_i - \bar{y})^2 = S_{yy}$
- Regression sum of squares: $SSR = \Sigma(\hat{y}_i - \bar{y})^2 = S_{xy}^2/S_{xx}$
- Error sum of squares: $SSE = \Sigma(y_i - \hat{y}_i)^2 = S_{yy} - S_{xy}^2/S_{xx}$
- Regression identity: $SST = SSR + SSE$
- Coefficient of determination: $r^2 = \dfrac{SSR}{SST}$
- Linear correlation coefficient:

$$r = \frac{\frac{1}{n-1}\Sigma(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y} \quad \text{or} \quad r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Chapter 16 Analysis of Variance (ANOVA)

- Notation in one-way ANOVA:
  - k = number of populations
  - n = total number of observations
  - $\bar{x}$ = mean of all n observations
  - n_j = size of sample from Population j
  - $\bar{x}_j$ = mean of sample from Population j
  - s_j^2 = variance of sample from Population j
  - T_j = sum of sample data from Population j
- Defining formulas for sums of squares in one-way ANOVA:

$$SST = \Sigma(x_i - \bar{x})^2$$

$$SSTR = \Sigma n_j(\bar{x}_j - \bar{x})^2$$

$$SSE = \Sigma(n_j - 1)s_j^2$$

- One-way ANOVA identity: $SST = SSTR + SSE$
- Computing formulas for sums of squares in one-way ANOVA:

$$SST = \Sigma x_i^2 - (\Sigma x_i)^2/n$$

$$SSTR = \Sigma(T_j^2/n_j) - (\Sigma x_i)^2/n$$

$$SSE = SST - SSTR$$

- Mean squares in one-way ANOVA:

$$MSTR = \frac{SSTR}{k-1} \qquad MSE = \frac{SSE}{n-k}$$

- Test statistic for one-way ANOVA (independent samples, normal populations, and equal population standard deviations):

$$F = \frac{MSTR}{MSE}$$

with df = $(k - 1, n - k)$.

- **Confidence interval for $\mu_i - \mu_j$ in the Tukey multiple-comparison method (independent samples, normal populations, and equal population standard deviations):**

$$(\bar{x}_i - \bar{x}_j) \pm \frac{q_\alpha}{\sqrt{2}} \cdot s\sqrt{(1/n_i) + (1/n_j)}$$

where $s = \sqrt{MSE}$ and q_α is obtained for a q-curve with parameters k and $n - k$.

- **Test statistic for a Kruskal–Wallis test (independent samples, same-shape populations, all sample sizes 5 or greater):**

$$H = \frac{SSTR}{SST/(n-1)} \quad \text{or} \quad H = \frac{12}{n(n+1)}\sum_{j=1}^{k}\frac{R_j^2}{n_j} - 3(n+1)$$

where $SSTR$ and SST are computed for the ranks of the data, and R_j denotes the sum of the ranks for the sample data from Population j. H has approximately a chi-square distribution with df = $k - 1$.

Additional Formulas:

* **Sample Mean:**

$\bar{x} = \frac{\Sigma x}{n}$ or $\bar{x} = \frac{\Sigma fx}{n}$ where $n = \Sigma f$

* **Sample Variance:**

$$s^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1} \quad \text{or} \quad s^2 = \frac{\Sigma fx^2 - \frac{(\Sigma fx)^2}{n}}{n-1}$$

* **Sample Standard Deviation:** $s = \sqrt{s^2}$

* **Empirical Rule:**

Approximately 68% of the data is within 1 std. dev. of the mean.
Approximately 95% of the data is within 2 std. dev. of the mean.
Approximately 99.7% of the data is within 3 std. dev. of the mean.

* **Chebyshev's rule:**

at least $1 - \frac{1}{k^2}$ of the data is within k std. dev. of the mean.

* **sample z-score for x:** $Z = \frac{x - \bar{x}}{s}$