FREQUENTLY USED FORMULAS

n = sample size N = population size f = frequency

Chapter 2

Class Width $= \frac{\text{high} - \text{low}}{\text{number classes}}$ (increase to next integer)

Class Midpoint $= \frac{\text{upper limit} + \text{lower limit}}{2}$

Lower boundary = lower boundary of previous class + class width

Chapter 3

Sample mean $\bar{x} = \frac{\Sigma x}{n}$

Population mean $\mu = \frac{\Sigma x}{N}$

Weighted average $= \frac{\Sigma xw}{\Sigma w}$

Range = largest data value − smallest data value

Sample standard deviation $s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$

Computation formula $s = \sqrt{\frac{SS_x}{n - 1}}$ where
$SS_x = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$

Population standard deviation $\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}}$

Sample variance s^2

Population variance σ^2

Sample Coefficient of Variation $CV = \frac{s}{\bar{x}} \cdot 100$

Sample mean for grouped data $\bar{x} = \frac{\Sigma xf}{n}$

Sample standard deviation for grouped data
$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{\Sigma x^2 f - (\Sigma xf)^2/n}{n - 1}}$

Chapter 4

Probability of the complement of event A
$P(not\ A) = 1 - P(A)$

Multiplication rule for independent events
$P(A\ and\ B) = P(A) \cdot P(B)$

General multiplication rules
$P(A\ and\ B) = P(A) \cdot P(B, given\ A)$
$P(A\ and\ B) = P(B) \cdot P(A, given\ B)$

Addition rule for mutually exclusive events
$P(A\ or\ B) = P(A) + P(B)$

General addition rule
$P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$

Permutation rule $P_{n,r} = \frac{n!}{(n - r)!}$

Combination rule $C_{n,r} = \frac{n!}{r!(n - r)!}$

Chapter 5

Mean of a discrete probability distribution $\mu = \Sigma x P(x)$

Standard deviation of a discrete probability distribution
$\sigma = \sqrt{\Sigma(x - \mu)^2 P(x)}$

Given $L - a + bx$

$\mu_L = a + b\mu$ $\sigma_L = |b|\sigma$

Given $W = ax_1 + bx_2$ (x_1 and x_2 independent)

$\mu_W = a\mu_1 + b\mu_2$

$\sigma_W = \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2}$

For Binomial Distributions

r = number of successes; p = probability of success; $q = 1 - p$

Binomial probability distribution $P(r) = C_{n,r}p^r q^{n-r}$

Mean $\mu = np$

Standard deviation $\sigma = \sqrt{npq}$

Geometric Probability Distribution

n = number of trial on which first success occurs
$P(n) = p(1 - p)^{n-1}$

Poisson Probability Distribution

λ = mean number of successes over given interval
$P(\lambda) = \frac{e^{-\lambda}\lambda^r}{r!}$

Chapter 6

Raw score $x = z\sigma + \mu$ Standard score $z = \frac{x - \mu}{\sigma}$

Chapter 7

Mean of $\bar{x}$ distribution $\mu_{\bar{x}} = \mu$

Standard deviation of $\bar{x}$ distribution $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

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

Mean of $\hat{p}$ distribution $\mu_{\hat{p}} = p$

Standard deviation of $\hat{p}$ distribution $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$; $q = 1 - p$

Chapter 8

Confidence Interval

for $\mu (n \geq 30)$

$$\bar{x} - z_c \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_c \frac{\sigma}{\sqrt{n}}$$

for $\mu (n < 30)$

$$d.f. = n - 1$$

$$\bar{x} - t_c \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_c \frac{s}{\sqrt{n}}$$

for $p (np > 5$ and $nq > 5)$

$$\hat{p} - z_c \sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}} < p < \hat{p} + z_c \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

where $\hat{p} = r/n$

for difference of means ($n_1 \geq 30$ and $n_2 \geq 30$)

$$(\bar{x}_1 - \bar{x}_2) - z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

for difference of means ($n_1 < 30$ and/or $n_2 < 30$ and $\sigma_1 \approx \sigma_2$)

$$d.f. = n_1 + n_2 - 2$$

$$(\bar{x}_1 - \bar{x}_2) - t_c s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_c s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where $s = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$

for difference of proportions

where $\hat{p}_1 = r_1/n_1$; $\hat{p}_2 = r_2/n_2$; $\hat{q}_1 = 1 - \hat{p}_1$; $\hat{q}_2 = 1 - \hat{p}_2$

$$(\hat{p}_1 - \hat{p}_2) - z_c \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_c \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Sample Size for Estimating

means $n = \left(\dfrac{z_c \sigma}{E}\right)^2$

proportions

$n = p(1 - p)\left(\dfrac{z_c}{E}\right)^2$ with preliminary estimate for p

$n = \dfrac{1}{4}\left(\dfrac{z_c}{E}\right)^2$ without preliminary estimate for p

Chapter 9

Sample Test Statistics for Tests of Hypotheses

for $\mu (n \geq 30)$ $\quad z = \dfrac{\bar{x} - \mu}{\sigma\sqrt{n}}$

for $\mu (n < 30)$; $d.f. = n - 1$ $\quad t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

for p $\quad z = \dfrac{\hat{p} - p}{\sqrt{pq/n}}$ where $q = 1 - p$

for paired differences d $\quad t = \dfrac{\bar{d} - \mu_d}{s_d/\sqrt{n}}$ with $d.f. = n - 1$

difference of means, large sample

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

difference of means, small sample with $\sigma_1 \approx \sigma_2$;

$$d.f. = n_1 + n_2 - 2$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

where $s = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$

difference of proportions

$z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\hat{p}\hat{q}}{n_1} + \dfrac{\hat{p}\hat{q}}{n_2}}}$ where $\hat{p} = \dfrac{r_1 + r_2}{n_1 + n_2}$; $\hat{q} = 1 - \hat{p}$;

$\hat{p}_1 = r_1/n_1$; $\hat{p}_2 = r_2/n_2$

Chapter 10

Regression and Correlation

In all these formulas $SS_x = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$,

$SS_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$, $SS_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}$

Least squares line $y = a + bx$ where $b = \frac{SS_{xy}}{SS_x}$ and

$a = \bar{y} - b\bar{x}$

Standard error of estimate $S_e = \sqrt{\frac{SS_y - bSS_{xy}}{n-2}}$

where $b = \frac{SS_{xy}}{SS_x}$

Pearson product moment correlation coefficient

$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}}$

Coefficient of determination $= r^2$

Confidence interval for y

$y_p - E < y < y_p + E$ where y_p is the predicted y value for x

$E = t_c S_e \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{SS_x}}$ with $d.f. = n - 2$

Sample test statistic for r

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ with $d.f. = n - 2$

Sample test statistic test for b

$t = \frac{b - \beta}{S_e/\sqrt{SS_x}}$ with $d.f. = n - 2$

Confidence interval for β

$b - t_c \frac{S_e}{\sqrt{SS_x}} < \beta < b + t_c \frac{S_e}{\sqrt{SS_x}}$ with $d.f. = n - 2$

Chapter 11

$x_2 = \sum \frac{(O - E)^2}{E}$ where $E = \frac{\text{(row total)(column total)}}{\text{sample size}}$

Tests of Independence $d.f. = (R - 1)(C - 1)$

Goodness of fit $d.f. =$ (number of entries) $- 1$

Confidence Interval for σ^2; $d.f. = n - 1$

$\frac{(n-1)s^2}{\chi^2_U} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$

Sample test statistic for H_0: $\sigma^2 = k$; $d.f. = n - 1$

$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

Testing Two Variances

Sample test statistic $F = \frac{s_1^2}{s_2^2}$

where $s_1^2 \geq s_2^2$

$d.f._N = n_1 - 1$; $d.f._D = n_2 - 1$

ANOVA

k = number of groups; N = total sample size

$SS_{TOT} = \Sigma x^2_{TOT} - \frac{(\Sigma x_{TOT})^2}{N}$

$SS_{BET} = \sum_{\text{all groups}} \left(\frac{(\Sigma x_i)^2}{n_i}\right) - \frac{(\Sigma x_{TOT})^2}{N}$

$SS_W = \sum_{\text{all groups}} \left(\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n_i}\right)$

$SS_{TOT} = SS_{BET} + SS_W$

$MS_{BET} = \frac{SS_{BET}}{d.f._{BET}}$ where $d.f._{BET} = k - 1$

$MS_W = \frac{SS_W}{d.f._W}$ where $d.f._W = N - k$

$F = \frac{MS_{BET}}{MS_W}$ where $d.f.$ numerator $= d.f._{BET} = k - 1$;

$d.f.$ denominator $= d.f._W = N - k$

Two-Way ANOVA

r = number of rows; c = number of columns

Row factor F: $\frac{MS \text{ row factor}}{MS \text{ error}}$

Column factor F: $\frac{MS \text{ column factor}}{MS \text{ error}}$

Interaction F: $\frac{MS \text{ interaction}}{MS \text{ error}}$

with degrees of freedom for

row factor $= r - 1$ interaction $= (r - 1)(c - 1)$

column factor $= c - 1$ error $= rc(n - 1)$

Chapter 12

Sample test statistic for x = proportion of plus signs to all signs ($n \geq 12$)

$z = \frac{x - 0.5}{\sqrt{0.25/n}}$

Sample test statistic for R = sum of ranks

$z = \frac{R - \mu_R}{\sigma_R}$ where $\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$ and

$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$

Spearman rank correlation coefficient

$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$ where $d = x - y$