FREQUENTLY USED FORMULAS

n = sample size $\quad N$ = population size $\quad f$ = frequency

Chapter 2

Class width $= \dfrac{\text{high} - \text{low}}{\text{number of classes}}$ (increase to next integer)

Class midpoint $= \dfrac{\text{upper limit} + \text{lower limit}}{2}$

Lower boundary = lower boundary of previous class
+ class width

Chapter 3

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

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

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

Range = largest data value − smallest data value

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

Computation formula $s = \sqrt{\dfrac{\Sigma x^2 - (\Sigma x)^2/n}{n - 1}}$

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

Sample variance s^2

Population variance σ^2

Sample coefficient of variation $CV = \dfrac{s}{\bar{x}} \cdot 100$

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

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
$P(A^c) = 1 - P(A)$

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

General multiplication rules
$P(A \text{ and } B) = P(A) \cdot P(B|A)$
$P(A \text{ and } B) = P(B) \cdot P(A|B)$

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

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

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

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

Chapter 5

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

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

r = number of successes

λ = mean number of successes over given interval

$P(r) = \dfrac{e^{-\lambda}\lambda^r}{r!}$

Chapter 6

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

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

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

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

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

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

Chapter 7

Confidence Interval

for μ

$$\bar{x} - E < \mu < \bar{x} + E$$

where $E = z_c \dfrac{\sigma}{\sqrt{n}}$ when σ is known

$E = t_c \dfrac{s}{\sqrt{n}}$ when σ is unknown

with $d.f. = n - 1$

for p ($np > 5$ and $n(1 - p) > 5$)

$$\hat{p} - E < p < \hat{p} + E$$

where $E = z_c \sqrt{\dfrac{\hat{p}(1 - \hat{p})}{n}}$

$\hat{p} = \dfrac{r}{n}$

for $\mu_1 - \mu_2$ (independent samples)

$$(\bar{x}_1 - \bar{x}_2) - E < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + E$$

where $E = z_c \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ when σ_1 and σ_2 are known

$E = t_c \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$ when σ_1 or σ_2 is unknown

with $d.f.$ = smaller of $n_1 - 1$ and $n_2 - 1$

(*Note*: Software uses Satterthwaite's approximation for degrees of freedom $d.f.$)

for difference of proportions $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) - E < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + E$$

where $E = z_c \sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

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

$\hat{q}_1 = 1 - \hat{p}_1;\ \hat{q}_2 = 1 - \hat{p}_2$

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 8

Sample Test Statistics for Tests of Hypotheses

for μ (σ known) $\quad z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

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

for p ($np > 5$ and $nq > 5$) $\quad z = \dfrac{\hat{p} - p}{\sqrt{pq/n}}$

where $q = 1 - p$; $\hat{p} = r/n$

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

for difference of means, σ_1 and σ_2 known

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

for difference of means, σ_1 or σ_2 unknown

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

$d.f.$ = smaller of $n_1 - 1$ and $n_2 - 1$

(*Note*: Software uses Satterthwaite's approximation for degrees of freedom $d.f.$)

for difference of proportions

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{\bar{p}\,\bar{q}}{n_1} + \dfrac{\bar{p}\,\bar{q}}{n_2}}}$$

where $\bar{p} = \dfrac{r_1 + r_2}{n_1 + n_2}$ and $\bar{q} = 1 - \bar{p}$

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

Chapter 9

Regression and Correlation

Pearson product-moment correlation coefficient

$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

Least-squares line $\hat{y} = a + bx$

where $b = \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$

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

Coefficient of determination $= r^2$

Sample test statistic for r

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

Standard error of estimate $S_e = \sqrt{\dfrac{\Sigma y^2 - a\Sigma y - b\Sigma xy}{n - 2}}$

Confidence interval for y

$$\hat{y} - E < y < \hat{y} + E$$

where $E = t_c S_e \sqrt{1 + \dfrac{1}{n} + \dfrac{n(x - \bar{x})^2}{n\Sigma x^2 - (\Sigma x)^2}}$

with $d.f. = n - 2$

Useful Calculator Functions Syntax

- `1-Var Stats` - `1-VarStats DataList [, FreqList]`
- `binompdf` - `binompdf(`n`,` p`,` r`)`
- `binomcdf` - `binomcdf(`n`,` p`,` r`)`
- `normalcdf` - `normalcdf(lower bound, upper bound,` μ`,` σ`)`
- `invNorm` - `invNorm(`p`,` μ`,` σ`)`
- `invT` - `invT(`p`, degrees of freedom)`
- `ZInterval` - `ZInterval(`σ`,` $\bar{x}$`,` n`,` c`)` OR `ZInterval(`σ`, List, Freq,` c`)`
- `TInterval` - `TInterval(`$\bar{x}$`,` s_x`,` n`,` c`)` OR `TInterval(List, Freq,` c`)`
- `1-PropZInt` - `1-PropZInt(`x`,` n`,` c`)`
- `2-SampZInt` - `2-SampZInt(`σ_1`,` σ_2`,` $\bar{x}_1$`,` n_1`,` $\bar{x}_2$`,` n_2`,` c`)` OR `2-SampZInt(`σ_1`,` σ_2`, List1, List2, Freq1, Freq2,` c`)`
- `2-SampTInt` - `2-SampTInt(`$\bar{x}_1$`,` s_{x_1}`,` n_1`,` $\bar{x}_2$`,` s_{x_2}`,` n_2`,` c`, Pooled)` OR `2-SampTInt(List1, List2, Freq1, Freq2,` c`, Pooled)`
- `2-PropZInt` - `2-PropZInt(`x_1`,` n_1`,` x_2`,` n_2`,` c`)`
- `Z-Test` - `Z-Test(`μ_0`,` σ`,` $\bar{x}$`,` n`, alternate hypothesis)` OR `Z-Test(`μ_0`,` σ`, List, Freq, alternate hypothesis)`
- `T-Test` - `T-Test(`μ_0`,` $\bar{x}$`,` s`,` n`, alternate hypothesis)` OR `T-Test(`μ_0`, List, Freq, alternate hypothesis)`
- `1-PropZTest` - `1-PropZTest(`p_0`,` x`,` n`, alternate hypothesis)`
- `2-SampZTest` - `2-SampZTest(`σ_1`,` σ_2`,` $\bar{x}_1$`,` n_1`,` $\bar{x}_2$`,` n_2`, alternate hypothesis)` OR `2-SampZTest(`σ_1`,` σ_2`, List1, List2, Freq1, Freq2, alternate hypothesis)`
- `2-SampTTest` - `2-SampTTest(`$\bar{x}_1$`,` s_{x_1}`,` n_1`,` $\bar{x}_2$`,` s_{x_2}`,` n_2`, alternate hypothesis, Pooled)` OR `2-SampTTest(List1, List2, Freq1, Freq2, alternate hypothesis, Pooled)`
- `2-PropZTest` - `2-PropZTest(`x_1`,` n_1`,` x_2`,` n_2`, alternate hypothesis)`
- `LinReg(a+bx)` - `LinReg(a+bx) Xlist, Ylist`
- `LinRegTInt` - `LinRegTInt(Xlist, Ylist, Freq,` c`)`
- `LinRegTTest` - `LinRegTTest(Xlist, Ylist, Freq, alternate hypothesis)`
- χ^2`-Test` - χ^2`-Test (Observed matrix, Expected matrix)`
- χ^2`GOF-Test` - χ^2`GOF-Test (Observed list, Expected list, d.f)`
- χ^2`cdf` - χ^2`cdf(lower bound, upper bound, d.f.)`