STAT 1220
FORMULAS FOR ELEMENTS OF STATISTICS I

DESCRIPTIVE

Sample Standard Deviation $s = \sqrt{\dfrac{\sum x^2 - \left[\frac{(\sum x)^2}{n}\right]}{n-1}}$

PROBABILITY

Addition Rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

Conditional Probability Rule: $P(A|B) = \dfrac{P(A \text{ and } B)}{P(B)}$

Special Multiplication Rule: $P(A \text{ and } B) = P(A)P(B)$ when A and B are independent

DISCRETE RANDOM VARIABLES

$$E(x) = \mu = \sum xP(x) \qquad \sigma = \sqrt{[\sum x^2 P(x)] - \mu^2}$$

BINOMIAL PROBABILITY

$$P(x) = \frac{n!}{x!(n-x)!}p^x q^{n-x} \qquad \mu = np \qquad \sigma = \sqrt{npq}$$

SAMPLING DISTRIBUTIONS

$$\mu_{\bar{x}} = \mu \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Inference about a single population mean or proportion

INFERENCE CONDITIONS	CONFIDENCE INTERVAL	TEST STATISTIC
σ unknown, normal population	$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$, with $df = n-1$	$t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$
Inference about a population proportion	$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$

Sample size formulas

for population mean: $n = \dfrac{\left(z_{\alpha/2}\right)^2 \sigma^2}{E^2}$; for population proportion: $n = \dfrac{\left(z_{\alpha/2}\right)^2 \hat{p}\hat{q}}{E^2}$ where E is a bound of a confidence interval

Inference about two population means

CONDITIONS	CONFIDENCE INTERVAL	TEST STATISTIC
independent samples, σ_1 and σ_2 unknown but assumed equal, populations normal	$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$ with $df = (n_1 + n_2 - 2)$ where $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$
paired differences, normal population of differences	$\bar{d} \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}}$ with $df = n - 1$	$t = \frac{\bar{d} - \mu_d}{s_D/\sqrt{n}}$

CORRELATION AND REGRESSION

$$\hat{y} = a + bx,\ b = \frac{SS_{xy}}{SS_{xx}},\ a = \bar{y} - b\bar{x},\ \text{and}$$

$$SS_{xx} = \sum x^2 - \left[\frac{\left(\sum x\right)^2}{n}\right],\quad SS_{xy} = \sum xy - \left[\frac{\left(\sum x\right)\left(\sum y\right)}{n}\right],\quad SS_{yy} = \sum y^2 - \left[\frac{\left(\sum y\right)^2}{n}\right]$$

$$SSE = SS_{yy} - bSS_{xy},\quad s_e = \sqrt{\frac{SSE}{n-2}},\quad s_b = \frac{s_e}{\sqrt{SS_{xx}}},$$

$$r^2 = b\frac{SS_{xy}}{SS_{yy}} = 1 - \frac{SSE}{SS_{yy}},\quad r = \frac{SS_{xy}}{\sqrt{SS_{xx}\cdot SS_{yy}}}$$

CONFIDENCE INTERVAL FOR B :

$$b \pm t_{\alpha/2}\frac{s_e}{\sqrt{SS_{xx}}} \quad \text{with } df = n - 2$$

TEST STATISTIC FOR B :

$$t = \frac{b - B}{s_e/\sqrt{SS_{xx}}}$$

CONFIDENCE INTERVAL FOR:

Mean value of y at $x = x_0$

$$\hat{y} \pm t_{\alpha/2} s_e\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_{xx}}} \qquad df = n-2$$

PREDICTION INTERVAL FOR

Individual new value of y at $x = x_0$

$$\hat{y} \pm t_{\alpha/2} s_e\sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{SS_{xx}}} \qquad df = n-2$$