COMM 215 Business Statistics
List of formula provided in the Final Examination

Chapter 3 Describing Data

Sample mean: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

Sample variance:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]$$

Sample standard deviation: $s = \sqrt{s^2}$

Z score: $z = \frac{x - \text{mean}}{\text{standard deviation}}$

Coefficient of variation: $\frac{\text{standard deviation}}{\text{mean}} \times 100$ %

Chapter 4 Probability

The rule of complement: $P(\bar{E}) = 1 - P(E)$

The addition rule for two events:

$$P(E_1 or E_2) = P(E_1) + P(E_2) - P(\text{E}_1 and E_2)$$

Conditional probability: $P(E_1 | E_2) = \frac{P(\text{E}_1 and E_2)}{P(E_2)}$

The general multiplication rule:

$$P(\text{E}_1 and E_2) = P(E_2)P(E_1 | E_2)$$

Chapter 5 Discrete Probability Distributions

Mean (expected value) of a discrete random variable

$$E(x) = \sum xP(x)$$

Variance and standard deviation of a discrete random variable

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

Binomial probability formula

$$p(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

Mean, variance, and standard deviation of a binomial random variable

$$\mu_x = np,\ \sigma_x^2 = npq, \text{ and } \sigma_x = \sqrt{npq}$$

Chapter 6 Continuous Probability Distribution

Standard normal random variable: $z = \frac{x - \mu}{\sigma}$

Chapter 7 Sampling distribution

$$\mu_{\bar{x}} = \mu$$

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

Sample proportion: $\bar{p} = \frac{x}{n}$

$$\mu_{\bar{p}} = p$$

Standard error of $\bar{p}$:

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Chapter 8 Estimating Single Population Parameters

A z-based confidence interval for a population mean μ with σ known:

Margin of error: $e = z\frac{\sigma}{\sqrt{n}}$

Confidence interval = $\bar{x} \pm z\frac{\sigma}{\sqrt{n}}$

A t-based confidence interval for a population mean μ with σ unknown: $\bar{x} \pm t\frac{s}{\sqrt{n}}$

Confidence interval for the proportion: $\bar{p} \pm z\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

Chapter 9 Hypothesis Testing

$$z-Test\ for\quad mean\quad z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$t-Test\ for\quad mean\quad t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$z-Test\quad for\quad proportion\ z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Chapter 13 Goodness-of-Fit Tests

Chi-square goodness-of-fit test statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

Chi-square contingency test statistic

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \text{ with } df = (r-1)(c-1)$$

Chapter 14 Simple Linear Regression and Correlation Analysis

Sample correlation coefficient

$$r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\left[\sum (x-\bar{x})^2\right]\left[\sum (y-\bar{y})^2\right]}} = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n(\sum x^2) - (\sum x)^2\right]\left[n(\sum y^2) - (\sum y)^2\right]}}$$

Simple linear regression model: $y = \beta_0 + \beta_1 x + \varepsilon$

Least squares point estimates

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

Sum of squared residuals (Sum of squares error)

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum y^2 - b_0 \sum y - b_1 \sum xy$$

Total sum of squares: $SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$

Sum of squares regression: $SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$

Simple regression estimator for the standard error of the estimate: $s_\varepsilon = \sqrt{\frac{SSE}{n-2}}$

Coefficient of determination: $R^2 = r^2 = \frac{SSR}{SST}$

F test for the simple linear regression model:

$$F = \frac{\frac{SSR}{1}}{\frac{SSE}{(n-2)}} \qquad df = (D_1 = 1, D_2 = n-2)$$

Simple regression estimator for the standard error of the slope: $s_{b_1} = \frac{s_\varepsilon}{\sqrt{\sum (x-\bar{x})^2}} = \frac{s_\varepsilon}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}}}$

Test of hypothesis for slope: $t = \frac{b_1 - \beta_1}{s_{b_1}} \qquad df = n-2$

Confidence interval for $E(y) \mid x_p$:

$$\hat{y} \pm t s_\varepsilon \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x-\bar{x})^2}}$$

Prediction interval for $y \mid x_p$: $\hat{y} \pm t s_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x-\bar{x})^2}}$

Chapter 15 Multiple regression

The multiple regression model:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$

Standard error: $s_\varepsilon = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{MSE}$

Multiple coefficient of determination: $R^2 = \frac{SSR}{SST}$

An *F* test for the linear regression model:

$$F = \frac{\frac{SSR}{k}}{\frac{SSE}{n-k-1}} \qquad df = (D_1 = k, D_2 = n-k-1)$$

Testing the significance of each regression coefficient:

$$t = \frac{b_j - 0}{s_{b_j}} \qquad df = n-k-1$$