# COMM 215 Business Statistics
# List of formulae provided in the Final Examination

## Chapter 2 Descriptive Statistics

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

Sample variance:

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

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

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

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

## Chapter 3 Probability

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

The addition rule for two events:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Conditional probability: $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

The general multiplication rule:

$$P(A \cap B) = P(A)P(B \mid A)$$

## Chapter 4 Discrete random variables

Mean (expected value) of a discrete random variable

$$\mu_x = \sum_{\text{All } x} x p(x)$$

Variance and standard deviation of a discrete random variable

$$\sigma_x^2 = \sum_{\text{All } x} (x - \mu_x)^2 \, p(x) \qquad \sigma_x = \sqrt{\sigma_x^2}$$

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 5 Continuous random variables

The normal probability curve:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

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

## Chapter 6 Sampling distribution

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

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

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

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

## Chapter 7 Confidence intervals

A *z*-based confidence interval for a population mean $\mu$ with $\sigma$ known

$$\left[\bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right] = \left[\bar{x} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right), \bar{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right]$$

A *t*-based confidence interval for a population mean $\mu$ with $\sigma$ unknown

$$\left[\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)\right]$$

## Chapter 14 Chi Square test

A goodness of fit test for multinomial probabilities

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

A chi-square test for independence

$$\chi^2 = \sum_{\text{all cells}} \frac{\left(f_{ij} - \hat{E}_{ij}\right)^2}{\hat{E}_{ij}}$$

**Chapter 11 Correlation coefficient and simple linear regression analysis**

$$s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$= \frac{\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)/n}{n-1}$$

$s_x, s_y$ *are sample standard deviations*

$$SS_{xy} = (n-1)s_{xy}$$

$$SS_{xx} = (n-1)s_x^2 \qquad SS_{yy} = (n-1)s_y^2$$

Simple correlation coefficient

$$r = \frac{s_{xy}}{s_x s_y}$$

Simple linear regression model

$$y = \mu_{y|x} + \varepsilon = \beta_0 + \beta_1 x + \varepsilon$$

Least squares point estimates

$$b_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{s_{xy}}{s_x^2}$$

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

Sum of squared residuals

$$SSE = SS_{yy} - \frac{SS_{xy}^2}{SS_{xx}}$$

$$= \sum_{i=1}^{n} y_i^2 - b_0 \sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i y_i$$

Standard error

$$s = \sqrt{\frac{SSE}{n-2}}$$

Total variation

$$\sum (y_i - \bar{y})^2 = SS_{yy}$$

Unexplained variation

$$\sum (y_i - \hat{y}_i)^2 = SSE$$

Explained variation

$$\sum (\hat{y}_i - \bar{y})^2$$

Mean square error

$$s^2 = \frac{SSE}{n-2}$$

Simple coefficient of determination

$$r^2 = \frac{\text{explained variation}}{\text{total variation}}$$

F test for the simple linear regression model

$$\frac{\text{explained variation}}{\text{unexplained variation}/(n-2)}$$

(Estimated) Standard error of the estimator $b_1$

$$\frac{s}{\sqrt{SS_{xx}}}$$

$$\text{Distance value} = \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_{xx}}$$

(Estimated) Standard error of $\hat{y}$

$$s_{\hat{y}} = s\sqrt{\text{distance value}}$$

Confidence interval for a mean value of *y*

$$\left[\hat{y} \pm t_{\alpha/2} s\sqrt{\text{distance value}}\right]$$

Prediction interval for an individual value of *y*

$$\left[\hat{y} \pm t_{\alpha/2} s\sqrt{1 + \text{distance value}}\right]$$

**Chapter 12 Multiple regression**

The multiple regression model

$$y = \mu_{y|x_1, x_2, \cdots, x_k} + \varepsilon$$

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

Standard error

$$s = \sqrt{\frac{SSE}{n-(k+1)}}$$

Mean square error

$$s^2 = \frac{SSE}{n-(k+1)}$$

Multiple coefficient of determination

$$R^2 = \frac{\text{explained variation}}{\text{total variation}}$$

Multiple correlation coefficient $R = \sqrt{R^2}$

An *F* test for the linear regression model

$$\frac{\text{explained variation}/k}{\text{unexplained variation}/[n-(k+1)]}$$

Testing the significance of an independent variable

$$t = \frac{b_j}{s_{b_j}}$$