PEARSON

Formula Card and Guide to Descriptive and Inferential Techniques for *Analyzing Data and Making Decisions*

Second Edition, by Judith Skuce

CHAPTER 2

Calculating class width (p. 35)

$$\text{Class width} = \frac{\text{maximum value} - \text{minimum value}}{\sqrt{n}}$$

CHAPTER 3

Calculating the sample mean (p. 105)

$$\bar{x} = \frac{\sum x}{n}$$

Calculating the location of the median (p. 109)

$$0.5(n+1)$$

Calculating the sample standard deviation with the computational formula (p. 119)

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$$

Calculating the interquartile range (p. 129)

$$\text{Location of } Q_1 = 0.25(n+1)$$
$$\text{Location of } Q_3 = 0.73(n+1)$$
$$IQR = Q_3 - Q_1$$

CHAPTER 4

If two events are independent, then (p. 162)

$$P(A|B) = P(A) \text{ and}$$
$$P(B|A) = P(B)$$

Calculating the conditional probability of A given B (p. 166)

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

Calculating the probability of A and B (p. 167)

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

Calculating the probability of A and B when A and B are independent events (p. 169)

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

Calculating the probability of A or B (p. 170)

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

Calculating the probability of A or B when A and B are mutually exclusive (p. 171)

$$P(A \text{ or } B) = P(A) + P(B)$$

Calculating the probability of an event A using the complement rule (p. 172)

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

CHAPTER 5

Calculating the mean of a discrete probability distribution (p. 187)

$$\mu = \sum(x \cdot P(x))$$

Calculating the standard deviation of a discrete probability distribution using the computational formula (p. 187)

$$\sigma = \sqrt{\sum x^2 P(x) - \mu^2}$$

Calculating the mean and standard deviation of a binomial random variable (p. 190)

$$\mu = np$$
$$\sigma = \sqrt{npq}$$

Calculating the probability of x successes in binomial experiment with n trials and probability of success p (p. 194)

$$P(x \text{ successes}) = \binom{n}{x} p^x q^{n-x}$$

Calculating the z-score to standardize a normal random variable (p. 212)

$$z = \frac{x - \mu}{\sigma}$$

Calculating the x-value that corresponds to a particular normal probability (p. 214)

$$x = \mu + z \cdot \sigma$$

CHAPTER 6

Calculating the mean and standard error of the sampling distribution of sample means (p. 233)

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

Calculating the mean and standard error of the sampling distribution of sample proportions (p. 248)

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$
$$\mu_{\hat{p}} = p$$

Conditions required for the normality of the sampling distribution of the sample proportions (p. 248)

$$np \geq 10 \text{ and } nq \geq 10$$

CHAPTER 7

Test statistic for the hypothesis test of a proportion (p. 274)

$$z = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test statistic for the hypothesis test of the mean (p. 276)

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

CHAPTER 8

Constructing the confidence interval estimate for p (p. 304)

$$\hat{p} \pm (\text{critical } z\text{-score})\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Constructing the confidence interval estimate for μ (p. 309)

$$\bar{x} \pm (\text{critical } t\text{-score})\left(\frac{s}{\sqrt{n}}\right)$$

Calculating the sample size required for an estimate of μ (p. 314)

$$n = \left(\frac{(z\text{-score})(s)}{\text{HW}}\right)^2$$

Calculating sample size required for an estimate of p (p. 316)

$$n = \hat{p}\hat{q}\left(\frac{z\text{-score}}{\text{HW}}\right)^2$$

CHAPTER 9

Test statistic for the hypothesis test of the mean difference in matched pairs (p. 335)

$$t = \frac{\bar{x}_D - \mu_D}{s_D/\sqrt{n_D}}$$

Constructing a confidence interval estimate for the mean difference in matched pairs (p. 335)

$$\bar{x}_D \pm \text{critical } t\text{-score}\left(\frac{s_D}{\sqrt{n_D}}\right)$$

Test statistic for W, in the Wilcoxon Signed Rank Sum Test, $n \geq 25$ (p. 353)

$$z = \frac{W - \mu_W}{\sigma_W} = \frac{W - \left(\frac{n_W(n_W+1)}{4}\right)}{\sqrt{\frac{n_W(n_W+1)(2n_W+1)}{24}}}$$

Test statistic for the sign test (p. 361)

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n_{ST}}}}$$

CHAPTER 10

Test statistic for the hypothesis test of the difference in means, independent samples, unequal variances, p. 380

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Constructing a confidence interval estimate for the difference in population means, independent samples, unequal variances (p. 381)

$$(\bar{x}_1 - \bar{x}_2) \pm t\text{-score}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Test statistic for W_1, in the Wilcoxon Rank Sum Test, $n_i \geq 10$ (p. 391)

$$z = \frac{W_1 - \mu_{W_1}}{\sigma_{W_1}} = \frac{W_1 - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}}}$$

CHAPTER 11

Calculating the mean square for within-sample variation (p. 411)

$$\text{MS}_{\text{within}} = \frac{\text{SS}_{\text{within}}}{n_T - k} = \frac{\text{SS}_1 + \text{SS}_2 + \cdots + \text{SS}_k}{n_T - k}$$

Calculating the mean square for between-sample variation (p. 412)

$$\text{MS}_{\text{between}} = \frac{\text{SS}_{\text{between}}}{k-1} = \frac{n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \cdots + n_k(\bar{x}_k - \bar{\bar{x}})^2}{k-1}$$

Calculating the test statistic for one-factor ANOVA (p. 419)

$$F = \frac{\text{MS}_{\text{between}}}{\text{MS}_{\text{within}}}$$

Calculating the Tukey-Kramer confidence interval for $(\mu_i - \mu_j)$ (p. 423)

$$(\bar{x}_i - \bar{x}_j) \pm q\text{-score}\sqrt{\frac{\text{MS}_{\text{within}}}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

CHAPTER 12

Calculating the pooled proportion for the hypothesis test of two proportions, with proportions hypothesized to be equal (p. 439)

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Test statistic for the hypothesis test of two proportions, proportions hypothesized to be equal (p. 445)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - \mu_{\hat{p}_1 - \hat{p}_2}}{s_{\hat{p}_1 - \hat{p}_2}} = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Test statistic for the hypothesis test of two proportions, proportions hypothesized to differ by a fixed amount (non-zero) (p. 446)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - \mu_{\hat{p}_1 - \hat{p}_2}}{s_{\hat{p}_1 - \hat{p}_2}} = \frac{(\hat{p}_1 - \hat{p}_2) - \mu_{\hat{p}_1 - \hat{p}_2}}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

Constructing a confidence interval estimate for the difference in population proportions (p. 444)

$$(\hat{p}_1 - \hat{p}_2) \pm z\text{-score}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Calculating Chi-square test statistic (pp. 449, 461)

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

Calculating expected values for a contingency table (p. 460)

$$\text{expected value} = \frac{(\text{row total})(\text{column total})}{(\text{grand total})}$$

CHAPTER 13

Calculating the residual for a least-squares line (p. 489)

$$e_i = y_i - \hat{y}_i$$

Test statistic for a hypothesis test about the slope of the regression line (p. 509)

$$t = \frac{b_1}{s_{b_1}}$$

CHAPTER 14

Calculating the test statistic for hypothesis test of overall regression model (p. 542)

$$F = \frac{\frac{\text{SSR}}{k}}{\frac{\text{SSE}}{n - (k+1)}} = \frac{\text{MSR}}{\text{MSE}}$$

A Guide to the Descriptive and Inferential Techniques of *Analyzing Data and Making Decisions*

Type of Investigation	Type of Data	Graph	Numerical Measure(s)	Inferential Techniques
What are the characteristics of this distribution of quantitative data?	Single-variable quantitative data • Organize in a frequency distribution	• Histogram • Comment on shape and skewness	• Measure of central tendency and measure of variability • See Guide to Decision Making: Choosing a Measure of Central Tendency on page 114 • See Guide to Decision Making: Choosing a Measure of Variability on page 131	• Hypothesis tests about μ (Chapter 7) • Confidence intervals about μ (Chapter 8)
Compare two or more distributions of quantitative data	Two or more sets of single-variable quantitative data • Organize in frequency distributions (use relative frequencies where appropriate)	• Two or more histograms set up for easy comparison • Comment on shape and skewness • See Guide to Technique: Comparing Histograms on page 60	• Measures of central tendency and measures of variability • See Guide to Decision Making: Choosing a Measure of Central Tendency on page 114 • See Guide to Decision Making: Choosing a Measure of Variability on page 131	Matched Pairs • Hypothesis tests about μ_D (Chapter 9) • Confidence intervals about μ_D (Chapter 9) • Wilcoxon Signed Rank Sum Test (Chapter 9) Independent Samples • *t*-test of $\mu_1 - \mu_2$ (Chapter 10) • Confidence intervals about $\mu_1 - \mu_2$ (Chapter 10) • Wilcoxon Rank Sum Test (Chapter 10) • ANOVA (Chapter 11)
What are the characteristics of the qualitative categories?	Collection of data with one qualitative characteristic, with two or more levels • Organize in a simple table	Bar graph (or pie chart, if there are relatively few categories and share of the total is important)	Frequencies or relative frequencies of categories	• Hypothesis test about p (Chapter 7) • Confidence intervals about p (Chapter 8) • χ^2 hypothesis test about goodness of fit (Chapter 12)
Compare two or more distributions of qualitative data • What are the characteristics of the qualitative categories? • Are these qualitative categories related?	Collection of data with two (or more) qualitative characteristics, each with two or more levels • Organize in a contingency table (at least a 2×2 table)	Bar graph with two or more bars in each category	Frequencies or relative frequencies of categories	• Hypothesis test about difference in matched pairs of ranked data, Sign Test (Chapter 9) • Hypothesis test about locations of populations of independent ranked data, Wilcoxon Rank Sum Test (Chapter 10) • Hypothesis test about $p_1 - p_2$ (Chapter 12) • Confidence intervals about $p_1 - p_2$ (Chapter 12) • χ^2 hypothesis test for testing many proportions, or testing for independence (Chapter 12)
What are the trends over time? • Compare trends over time for two or more variables	Time-series data (generally quantitative)	• Line graph, with time on the *x*-axis • Multiple lines on the graph if time series are being compared	Absolute change or percentage change over time • Significant increases or decreases noted, citing specific values • Significant changes in trend noted, citing specific values	
Are these variables related/associated and, if so, how?	• Paired data, quantitative or ranked, of the form $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$	Scatter diagram for quantitative data • Explanatory variable on the *x*-axis • Response variable on the *y*-axis	• Pearson r or Spearman r • Coefficient of determination, R^2 • See Guide to Decision Making: Choosing a Measure of Association on page 142	• Hypothesis test about the regression relationship (Chapters 13 & 14) – Regression confidence and prediction intervals (Chapters 13 & 14)
	• Quantitative response (y) variable, with quantitative or qualitative explanatory variables (more than one)		• Adjusted R^2	