NOTATION  The following notation is used on this card:

- $n$ = sample size
- $\mu$ = population mean
- $\sigma$ = population stdev
- $\bar{x}$ = sample mean
- $s$ = sample stdev
- $Q_j$ = $j$th quartile
- $P$ = population proportion
- $O$ = observed frequency
- $E$ = expected frequency
- $d$ = paired difference
- $\hat{p}$ = sample proportion
- $S_{xx}$ = population standard deviation
- $S_{xy}$ = linear correlation coefficient
- $Q_0$ = first quartile
- $Q_3$ = third quartile
- $N$ = population size
- $\sigma$ = sample standard deviation
- $r$ = coefficient of determination
- $b_1$ = sample slope
- $b_0$ = sample intercept
- $\hat{y}$ = fitted value
- $\hat{\sigma}^2$ = estimated variance
- $\hat{\sigma}^2$ = estimated standard deviation
- $\bar{y}$ = sample mean
- $s^2$ = sample variance
- $SSR$ = regression sum of squares
- $SSE$ = error sum of squares
- $SST$ = total sum of squares
- $R^2$ = coefficient of determination
- $IQR$ = interquartile range
- $z$ = standard normal variable

CHAPTER 5 Probability and Random Variables

- Probability for equally likely outcomes:
  
  \[ P(E) = \frac{f}{N} \]
  
  where $f$ denotes the number of ways event $E$ can occur and $N$ denotes the total number of outcomes possible.

- Special addition rule:
  
  \[ P(A \text{ or } B \text{ or } C \cdots) = P(A) + P(B) + P(C) + \cdots \]
  
  (A, B, C, \ldots mutually exclusive)

- Complementation rule: $P(E) = 1 - P(\text{not } E)$

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

- Mean of a discrete random variable $X$:
  
  \[ \mu = \Sigma_x P(X = x) \]

- Standard deviation of a discrete random variable $X$:
  
  \[ \sigma = \sqrt{\Sigma(x - \mu)^2 P(X = x)} \] or \[ \sigma = \sqrt{\Sigma x^2 P(X = x) - \mu^2} \]

- Factorial: $k! = k(k - 1) \cdots 2 \cdot 1$

- Binomial coefficient:
  
  \[ \binom{n}{x} = \frac{n!}{x!(n-x)!} \]

- Binomial probability formula:
  
  \[ P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \]

  where $n$ denotes the number of trials and $p$ denotes the success probability.

- Mean of a binomial random variable: $\mu = np$

- Standard deviation of a binomial random variable: $\sigma = \sqrt{np(1-p)}$

CHAPTER 7 The Sampling Distribution of the Sample Mean

- Mean of the variable $\bar{x}$: $\mu_{\bar{x}} = \mu$

- Standard deviation of the variable $\bar{x}$: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

CHAPTER 8 Confidence Intervals for One Population Mean

- Sample size for estimating $\mu$:

  \[ n = \left( \frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 \]

  rounded up to the nearest whole number.


**CHAPTER 10 Inferences for Two Population Means**

- Studentized version of the variable \( \bar{x} \):
  \[
  t = \frac{\bar{x} - \mu}{s/\sqrt{n}}
  \]

- \( t \)-interval for \( \mu \) (\( \sigma \) unknown, normal population or large sample):
  \[
  \bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}
  \]
  with \( df = n - 1 \).

**CHAPTER 9 Hypothesis Tests for One Population Mean**

- \( z \)-test statistic for \( H_0: \mu = \mu_0 \) (\( \sigma \) known, normal population or large sample):
  \[
  z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}
  \]

- \( t \)-test statistic for \( H_0: \mu = \mu_0 \) (\( \sigma \) unknown, normal population or large sample):
  \[
  t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}
  \]
  with \( df = n - 1 \).

**CHAPTER 11 Inferences for Population Proportions**

- Sample proportion:
  \[
  \hat{p} = \frac{x}{n},
  \]
  where \( x \) denotes the number of members in the sample that have the specified attribute.

- One-sample \( z \)-interval for \( p \):
  \[
  \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}
  \]
  (Assumption: both \( x \) and \( n - x \) are 5 or greater)

- Margin of error for the estimate of \( p \):
  \[
  E = z_{\alpha/2} \cdot \sqrt{\hat{p}(1 - \hat{p})/n}
  \]

- Sample size estimating \( p \):
  \[
  n = \frac{0.25 \left( \frac{z_{\alpha/2}}{E} \right)^2}{\hat{p}(1 - \hat{p})}
  \]
  or
  \[
  n = \frac{\hat{p}(1 - \hat{p})}{\hat{p}(1 - \hat{p})}
  \]
  rounded up to the nearest whole number (\( g \) = “educated guess”)

- One-sample \( z \)-test statistic for \( H_0: p = p_0 \):
  \[
  z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}
  \]
  (Assumption: both \( np_0 \) and \( n(1 - p_0) \) are 5 or greater)

- Pooled sample proportion:
  \[
  \hat{p}_p = \frac{x_1 + x_2}{n_1 + n_2}
  \]

- Two-sample \( z \)-test statistic for \( H_0: p_1 = p_2 \):
  \[
  z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p(1 - \hat{p}_p)(1/n_1) + (1/n_2)}}
  \]
  (Assumptions: independent samples; \( x_1, n_1 - x_1, x_2, n_2 - x_2 \) are all 5 or greater)
CHAPTER 13 Analysis of Variance (ANOVA)

- Two-sample z-interval for \( p_1 - p_2 \):
  \[
  (\hat{p}_1 - \hat{p}_2) \pm z_{a/2} \cdot \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)/n_1 + \hat{p}_2(1 - \hat{p}_2)/n_2}{}}
  \]
  \( \text{(Assumptions: independent samples; } x_1, n_1 - x_1, x_2, n_2 - x_2 \text{ are all 5 or greater) } \)

- Margin of error for the estimate of \( p_1 - p_2 \):
  \[
  E = z_{a/2} \cdot \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)/n_1 + \hat{p}_2(1 - \hat{p}_2)/n_2}{}}
  \]

- Sample size for estimating \( p_1 - p_2 \):
  \[
  n_1 = n_2 = 0.5 \left( \frac{z_{a/2}^2}{E} \right)^2
  \]
  or
  \[
  n_1 = n_2 = \left( \frac{\hat{p}_{18}(1 - \hat{p}_{18}) + \hat{p}_{28}(1 - \hat{p}_{28})}{E} \right)^2
  \]
  rounded up to the nearest whole number (\( g = \text{“educated guess”} \))

CHAPTER 12 Chi-Square Procedures

- Expected frequencies for a chi-square goodness-of-fit test:
  \[
  E = np
  \]

- Test statistic for a chi-square goodness-of-fit test:
  \[
  \chi^2 = \Sigma(O - E)^2 / E
  \]
  with df = \( k - 1 \), where \( k \) is the number of possible values for the variable under consideration.

- Expected frequencies for a chi-square independence test:
  \[
  E = \frac{R \cdot C}{n}
  \]
  where \( R \) = row total and \( C \) = column total.

- Test statistic for a chi-square independence test:
  \[
  \chi^2 = \Sigma(O - E)^2 / E
  \]
  with df = \((r - 1)(c - 1)\), where \( r \) and \( c \) are the number of possible values for the two variables under consideration.

CHAPTER 13 Analysis of Variance (ANOVA)

- Notation in one-way ANOVA:
  - \( k \) = number of populations
  - \( n \) = total number of observations
  - \( \bar{x} \) = mean of all \( n \) observations
  - \( n_j \) = size of sample from Population \( j \)
  - \( \bar{x}_j \) = mean of sample from Population \( j \)
  - \( s_j^2 \) = variance of sample from Population \( j \)
  - \( T_j \) = sum of sample data from Population \( j \)

- Defining formulas for sums of squares in one-way ANOVA:
  \[
  \begin{align*}
  SST &= \Sigma(x - \bar{x})^2 \\
  SSTR &= \Sigma n_j(\bar{x}_j - \bar{x})^2 \\
  SSE &= \Sigma(n_j - 1)s_j^2
  \end{align*}
  \]

- One-way ANOVA identity: \( SST = SSTR + SSE \)

- Computing formulas for sums of squares in one-way ANOVA:
  \[
  \begin{align*}
  SST &= \Sigma(x - \bar{x})^2 / n \\
  SSTR &= \Sigma(T_j^2 / n_j) - (\Sigma x)^2 / n \\
  SSE &= SST - SSTR
  \end{align*}
  \]

- Mean squares in one-way ANOVA:
  \[
  MSTR = \frac{SSTR}{k - 1}, \quad MSE = \frac{SSE}{n - k}
  \]

- Test statistic for one-way ANOVA (independent samples, normal populations, and equal population standard deviations):
  \[
  F = \frac{MSTR}{MSE}
  \]
  with df = \((k - 1, n - k)\).

CHAPTER 14 Inferential Methods in Regression and Correlation

- Population regression equation: \( y = \beta_0 + \beta_1 x \)

- Standard error of the estimate: \( s_e = \sqrt{\frac{SSE}{n - 2}} \)

- Test statistic for \( H_0: \beta_1 = 0 \):
  \[
  t = \frac{b_1}{s_e / \sqrt{Sxx}}
  \]
  with df = \( n - 2 \).

- Confidence interval for \( \beta_1 \):
  \[
  b_1 \pm t_{a/2} \cdot \frac{s_e}{\sqrt{Sxx}}
  \]
  with df = \( n - 2 \).

- Confidence interval for the conditional mean of the response variable corresponding to \( x_p \):
  \[
  \hat{y}_p \pm t_{a/2} \cdot s_e \sqrt{\frac{1}{n} + \frac{(x_p - \Sigma x/n)^2}{Sxx}}
  \]
  with df = \( n - 2 \).

- Prediction interval for an observed value of the response variable corresponding to \( x_p \):
  \[
  \hat{y}_p \pm t_{a/2} \cdot s_e \sqrt{1 + \frac{1}{n} + \frac{(x_p - \Sigma x/n)^2}{Sxx}}
  \]
  with df = \( n - 2 \).

- Test statistic for \( H_0: \rho = 0 \):
  \[
  t = \frac{r}{\sqrt{1 - r^2}} \sqrt{\frac{n - 2}{n - 2}}
  \]
  with df = \( n - 2 \).