

Chapter 2 Descriptive Statistics

The sample mean, \bar{x}	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$
The sample variance, s^2	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}$
The sample standard deviation, s	$s = \sqrt{s^2}$
Computational formula for s^2	$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i \right)^2}{n} \right]$
z score	$z = \frac{x - \text{mean}}{\text{standard deviation}}$
The coefficient of variation	coefficient of variation = $\frac{\text{standard deviation}}{\text{mean}} \times 100$
The weighted mean	$\frac{\sum w_i x_i}{\sum w_i}$
Sample mean for grouped data	$\bar{x} = \frac{\sum f_i M_i}{\sum f_i} = \frac{\sum f_i M_i}{n}$
Sample variance for grouped data	$s^2 = \frac{\sum f_i (M_i - \bar{x})^2}{n-1}$
Population mean for grouped data	$\mu = \frac{\sum f_i M_i}{N}$
Population variance for grouped data	$\sigma^2 = \frac{\sum f_i (M_i - \mu)^2}{N}$

Chapter 3 Probability

Probabilities when all sample space outcomes are equally likely	$\frac{\text{the number of sample space outcomes that correspond to the event}}{\text{the total number of sample space outcomes}}$
The rule of complements	$P(\bar{A}) = 1 - P(A)$
The addition rule for two events	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

The addition rule for two mutually exclusive events	$P(A \cup B) = P(A) + P(B)$
The addition rule for N mutually exclusive events	$P(A_1 \cup A_2 \cup \dots \cup A_N) = P(A_1) + P(A_2) + \dots + P(A_N)$
Conditional probability	$P(A B) = \frac{P(A \cap B)}{P(B)}$ $P(B A) = \frac{P(A \cap B)}{P(A)}$
The general multiplication rule	$P(A \cap B) = P(A)P(B A)$ $= P(B)P(A B)$
Independent events	$P(A B) = P(A)$ $P(B A) = P(B)$
The multiplication rule for two independent events	$P(A \cap B) = P(A)P(B)$
The multiplication rule for N independent events	$P(A_1 \cap A_2 \cap \dots \cap A_N) = P(A_1)P(A_2) \dots P(A_N)$
Bayes' theorem	$P(S_i E) = \frac{P(S_i \cap E)}{P(E)} = \frac{P(S_i)P(E S_i)}{P(E)}$

Chapter 4 Discrete Random Variables

Mean (expected value) of a discrete random variable	$\mu_x = \sum_{\text{All } x} xp(x)$
Variance and standard deviation of a discrete random variable	$\sigma_x^2 = \sum_{\text{All } x} (x - \mu_x)^2 p(x)$ $\sigma_x = \sqrt{\sigma_x^2}$
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	$\mu_x = np, \sigma_x^2 = npq, \text{ and } \sigma_x = \sqrt{npq}$
Poisson probability formula	$P(x) = \frac{e^{-\mu} \mu^x}{x!}$
Mean, variance, and standard deviation of a Poisson random variable	$\mu_x = \mu, \sigma_x^2 = \mu, \sigma_x = \sqrt{\mu}$

Hypergeometric probability formula	$p(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$
Mean and variance of a hypergeometric random variable	$\mu = n \left(\frac{r}{N} \right) \text{ and } \sigma^2 = n \left(\frac{r}{N} \right) \left(\frac{N-r}{N} \right) \left(\frac{N-n}{N-1} \right)$

Chapter 5 Continuous Random Variables

The uniform probability curve	$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$
Mean and standard deviation of a uniform distribution	$\mu_x = \frac{a+b}{2} \text{ and } \sigma_x = \frac{b-a}{\sqrt{12}}$
The normal probability curve	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
z values	$z = \frac{x-\mu}{\sigma}$
The exponential probability curve	$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$
Mean and standard deviation of an exponential distribution	$\mu_x = \frac{1}{\lambda} \text{ and } \sigma_x = \frac{1}{\lambda}$

Chapter 7 Confidence Intervals

A z-based confidence interval for a population mean μ with σ 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 μ with σ unknown	$\left[\bar{x} \pm t_{\alpha/2} \left(\frac{s}{\sqrt{n}} \right) \right]$
Sample size when estimating μ	$n = \left(\frac{z_{\alpha/2} \sigma}{E} \right)^2$
A large-sample confidence interval for a population proportion p	$\left[\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$
Sample size when estimating p	$n = p(1-p) \left(\frac{z_{\alpha/2}}{E} \right)^2$
z-based confidence interval for $\mu_1 - \mu_2$	$\left[\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right]$

t -based confidence interval for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$	$\left[\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right]$
t -based confidence intervals for $\mu_1 - \mu_2$	$\left[\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right]$
Confidence interval for μ_d	$\left[\bar{d} \pm t_{\alpha/2} \left(\frac{s_d}{\sqrt{n}} \right) \right]$
Large-sample confidence interval for $p_1 - p_2$	$\left[\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \right]$

Chapter 9 Statistical Inferences Based on Two Samples

z test about $\mu_1 - \mu_2$	$z = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$
t test about $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$	$t = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$
t test about $\mu_1 - \mu_2$ when $\sigma_1^2 \neq \sigma_2^2$	$t = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

Chapter 10 Experimental Design and Analysis of Variance

One-way ANOVA sums of squares: between-groups sum of squares	$SSB = \sum_{i=1}^p n_i (\bar{x}_i - \bar{x})^2$
One-way ANOVA sums of squares: error sum of squares	$SSE = \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2 + \dots + \sum_{j=1}^{n_p} (x_{pj} - \bar{x}_p)^2$
One-way ANOVA sums of squares: total sum of squares	$SST = SSB + SSE$
One-way ANOVA F test	$F = \frac{MSB}{MSE} = \frac{SSB / (p-1)}{SSE / (n-p)}$
Randomized block sums of squares	$SST = SSB + SSBL + SSE, \text{ where}$ $SSB = b \sum_{i=1}^p (\bar{x}_{i.} - \bar{x})^2$

	$SSBL = p \sum_{j=1}^b (\bar{x}_{.j} - \bar{x})^2$ $SST = \sum_{i=1}^p \sum_{j=1}^b (x_{ij} - \bar{x})^2$ $SSE = SST - SSB - SSBL$
Two-way ANOVA sums of squares	$SST = SS(1) + SS(2) + SS(\text{int}) + SSE, \text{ where}$ $SST = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^m (x_{ij,k} - \bar{x})^2$ $SS(1) = bm \sum_{i=1}^a (\bar{x}_{i.} - \bar{x})^2$ $SS(2) = am \sum_{j=1}^b (\bar{x}_{.j} - \bar{x})^2$ $SS(\text{int}) = m \sum_{i=1}^a \sum_{j=1}^b (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$ $SSE = SST - SS(1) - SS(2) - SS(\text{int})$

Chapter 11 Correlation Coefficient and Simple Linear Regression Analysis

Simple correlation coefficient	$r = \frac{s_{xy}}{s_x s_y}$
Testing the significance of the population correlation coefficient	$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$
Simple linear regression model	$y = \mu_{y x} + \varepsilon = \beta_0 + \beta_1 x + \varepsilon$
Mean square error	$s^2 = \frac{SSE}{n-2}$
Standard error	$s = \sqrt{\frac{SSE}{n-2}}$
Least squares line (prediction equation)	$\hat{y} = b_0 + b_1 x$
Sum of squared residuals	$SSE = SS_{yy} - \frac{SS_{xy}^2}{SS_{xx}}$
The residual	$e = y - \hat{y}$
Standard error of the estimate b_1	$\frac{s_{b_1}}{b_1 - \beta_1}$
Testing the significance of the slope	$t = \frac{b_1}{s_{b_1}}$

Confidence interval for the slope	$[b_1 + t_{\alpha/2} s_{b_1}]$
Testing the significance of the y intercept	$t = \frac{b_0}{s_{b_0}}$, where $S_{b_0} = S \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{SS_{xx}}}$
Distance value	Distance value = $\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_{xx}}$
Standard error of \hat{y}	$s_{\hat{y}} = s \sqrt{\text{distance value}}$
Confidence interval for a mean value of y	$[\hat{y} \pm t_{\alpha/2} s \sqrt{\text{distance value}}]$
Prediction interval for an individual value of y	$[\hat{y} \pm t_{\alpha/2} s \sqrt{1 + \text{distance value}}]$
Unexplained variation	Unexplained variation = $\Sigma (y_i - \hat{y}_i)^2$
Explained variation	Explained variation = $\Sigma (\hat{y}_i - \bar{y})^2$
Total variation	Total variation = $\Sigma (y_i - \bar{y})^2$
Simple coefficient of determination	$r^2 = \frac{\text{explained variation}}{\text{total variation}}$
An F test for the simple linear regression model	$F(\text{model}) = \frac{\text{explained variation}}{(\text{unexplained variation}) / (n - 2)}$
Durbin-Watson test	$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$

Chapter 12 Multiple Regression

The multiple regression model	$y = \mu_{y x_1, x_2, \dots, x_k} + \varepsilon$ $= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$
Point prediction of an individual value of y	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$
Mean square error	$s^2 = \frac{SSE}{n - (k + 1)}$
Standard error	$s = \sqrt{\frac{SSE}{n - (k + 1)}}$
The least squares point estimates	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$
Point estimate of a mean value of y	$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$
Total variation	Total variation = $\Sigma (y_i - \bar{y})^2$

Explained variation	Explained variation = $\Sigma(\hat{y}_i - \bar{y})^2$
Unexplained variation	Unexplained variation = $\Sigma(y_i - \hat{y}_i)^2$
Multiple coefficient of determination	$R^2 = \frac{\text{explained variation}}{\text{total variation}}$
Multiple correlation coefficient	Multiple correlation coefficient = $R = \sqrt{R^2}$
Adjusted multiple coefficient of determination	$\bar{R}^2 = \left(R^2 - \frac{k}{n-1}\right) \left(\frac{n-1}{n-(k+1)}\right)$
An F test for the linear regression model	$F(\text{model}) = \frac{\text{explained variation} / k}{\text{unexplained variation} / [n-(k+1)]}$
Testing the significance of an independent variable	$t = \frac{b_j}{s_{b_j}}$
Confidence interval for β_j	$[b_j \pm t_{\alpha/2} s_{b_j}]$
Confidence interval for a mean value of y	$[\hat{y} \pm t_{\alpha/2} s \sqrt{\text{distance value}}]$
Prediction interval for an individual value of y	$[\hat{y} \pm t_{\alpha/2} s \sqrt{1 + \text{distance value}}]$
The quadratic regression model	$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$
The partial F test	$F = \frac{(SSE_R - SSE_C) / (k - g)}{SSE_C / [n - (k + 1)]}$

Chapter 13 Nonparametric Methods

Large-sample sign test	$z = \frac{S - 0.5 - 0.5n}{0.5\sqrt{n}}$
Kruskal-Wallis H statistic	$H = \frac{12}{n(n+1)} \sum_{i=1}^p \frac{T_i^2}{n_i} - 3(n+1)$
Spearman's rank correlation coefficient	$r = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\sqrt{\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n_1}\right) \left(\Sigma y^2 - \frac{(\Sigma y)^2}{n_2}\right)}}$

Chapter 14 Chi-Square Tests

A goodness of fit test for multinomial probabilities	$\chi^2 = \sum_{i=1}^k \frac{(f_i - E_i)^2}{E_i}$
--	---

A goodness of fit test for a normal distribution	$\chi^2 = \sum_{i=1}^k \frac{(f_i - E_i)^2}{E_i}$
A chi-square test for independence	$\chi^2 = \sum_{\text{all cells}} \frac{(f_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$

Chapter 15 Decision Theory

Expected value of perfect information	EVPI = expected payoff under certainty – expected payoff under risk
Expected value of sample information	EVSI = EPS – EPNS
Expected net gain of sampling	ENGSI = EVSI – cost of sampling

Chapter 16 Time Series Forecasting

No trend	$y_t = \beta_0 + \varepsilon_t$
Linear trend	$y_t = \beta_0 + \beta_1 t + \varepsilon_t$
Quadratic trend	$\beta_0 + \beta_1 t + \beta_2 t^2$
Simple exponential smoothing	$S_T = \alpha y_T + (1 - \alpha) S_{T-1}$
Double exponential smoothing	$y_t = \beta_0 + \beta_1 t + \varepsilon_t$
Mean absolute deviation (MAD)	$\text{MAD} = \frac{\sum \text{abs}(y_t - \hat{y}_t)}{t}$
Mean squared deviation (MSD)	$\text{MSD} = \frac{\sum (y_t - \hat{y}_t)^2}{t}$
Simple index	$\frac{y_t}{y_0} \times 100$
Aggregate price index	$\left(\frac{\sum p_t}{\sum p_0} \right) \times 100$
Laspeyres index	$\frac{\sum p_t q_0}{\sum p_0 q_0} \times 100$
Paasche index	$\frac{\sum p_t q_t}{\sum p_0 q_t} \times 100$

Chapter 17 Process Improvement Using Control Charts

Centre line and control limits for an \bar{x} chart	Centre line $\bar{x} = \bar{\bar{x}}$, $UCL_{\bar{x}} = \bar{\bar{x}} + A_2 \bar{R}$, $LCL_{\bar{x}} = \bar{\bar{x}} - A_2 \bar{R}$,
Centre line and control limits for an R chart	Centre line $R = \bar{R}$, $UCL_R = D_4 \bar{R}$, $LCL_R = D_3 \bar{R}$,
Natural tolerance limits for normally distributed process measurements	$\left[\bar{\bar{x}} \pm 3 \left(\frac{\bar{R}}{d_2} \right) \right] = \left[\bar{\bar{x}} - 3 \left(\frac{\bar{R}}{d_2} \right), \bar{\bar{x}} + 3 \left(\frac{\bar{R}}{d_2} \right) \right]$
Centre line and control limits for a p chart	Centre line $= \bar{p}$, $UCL = \bar{p} + 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$, $LCL = \bar{p} - 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$