Errata and Corrigenda for *Statistics for Chemical and Material Engineers: A Modern Approach*

Last Update: July 8th, 2020

Page, line	Current Form	Correction
p. 3, line 14	…some information about the most common value in the data set.	…some information about the **central or typical** value in the data set
p. 8, fourth line from Equation (1.18)	from a normal distribution and is not very robust.	from a normal distribution**. It** is not very robust.
p.8, Equation (1.19) and line before	σ_{mad}	$\hat{\sigma}_{rob}$ [1]
p. 16, line 5	…tail is above,…	…tail is above **the straight line**,…
p. 31, line 7 p. 32, line 4	$\Omega(\mathbb{S}, \mathbb{F}, P)$	$\Omega = (\mathbb{S}, \mathbb{F}, P)$

---

[1] This makes the symbols consistent and also emphasis that we are dealing with the robust estimate of the standard deviation.

Page, line	Current Form	Correction
p. 31, line 16, p. 32ff; Table 2.5, 2.6	probability function	probability measure function (footnote: Also called a *probability measure*, a *probability function*, or even just *probability*.)
p. 31, line 4	… if an event $\mathbb{E}$ is an element in $\mathbb{F}$, or $\mathbb{E} \in \mathbb{F}$, then the set of $\mathbb{F}$ excluding $\mathbb{E}$ is also an element of $\mathbb{F}$, or $\mathbb{F} \backslash \mathbb{E} \in \mathbb{F}$.	…if an event **$\mathbb{E} \subseteq \mathbb{S}$** is an element in $\mathbb{F}$, or $\mathbb{E} \in \mathbb{F}$, then the set of **$\mathbb{S}$** excluding $\mathbb{E}$ is also an element of $\mathbb{F}$, or **$\mathbb{S}$**$\backslash \mathbb{E} \in \mathbb{F}$.
p. 31, line 6	… that is, the union of countable many subsets of $\mathbb{F}$ is in $\mathbb{F}$.	… that is, the union of countable many **elements** of $\mathbb{F}$ is in $\mathbb{F}$.
p. 32, line 8	… from all possible combinations of the original set…	… from all possible combinations (subsets) of the original set…
p. 32, line 26 to p. 34, line 22 (all text between Examples 2.1 and 2.2)	Let X be a random variable…(2.10).	*See* **Appendix VII: Updating Text in §2.1** *for the updated text.*
p. 34, 10	written as m_n,	written as **m_i**,
p. 35, last Equation	$f(x) = \frac{3}{8}x^3$	$f(x) = \frac{3}{8}x^2$ should be *squared* rather than *cubed*.

Page, line	Current Form	Correction
p. 35, line 15	is a probability density function. In general,	In general, (*that text is unnecessary and has been removed.*)
p. 37, Equation (2.13)	$E(x)=\int_{-\infty}^{\infty} xf(x)dx=\mu.$	$E(X)=\int_{-\infty}^{\infty} xf(x)dx=\mu.$
p. 37, right after Equation (2.13)	*Adding text.*	For the discrete case, the expectation operator can be computed as $E(X)=\sum_{x\in\mathbb{S}} xP(X=x)=\mu. \qquad (2.13')$
p. 37, last equation of Example 2.4	$E(3XY)=3E(XY)$ $=3\big(E(X)E(Y)+\mathrm{cov}(X,Y)\big)=3\big((5)(2)-2\big)$ $=24$	$E(3XY)=3E(XY)$ $=3\big(E(X)E(Y)+\mathrm{cov}(X,Y)\big)=3\big((5)(2)+2\big)$ $=36$ It should be “+2” rather than “–2” in the second line of the equation. As well, this will then change the final answer.

Page, line	Current Form	Correction
p. 41 (last Equation on page)	$\mu_Z = \int_0^\infty z f_Z(z)\,dy = \int_0^\infty 4ze^{-4z}dy = 4\times\frac{1}{16} = 0.25$	$\mu_Z = \int_0^\infty z f_Z(z)\,dy = \int_0^\infty 4ze^{-4z}d\underline{z} = 4\times\frac{1}{16} = 0.25$ [2]
p. 43, Section 4.2, line 6	Except for the last three distributions which are discrete	Except for the last **two** distributions which are discrete
p. 44, Equation (2.26)	$\Phi(z) = P(X \le z) = \int_{-\infty}^{z} \frac{1}{\sigma^2\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2}} dx$	$\Phi(z) = P(X \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-x^2}{2}} dx$ The μ and σ have fixed values of 0 and 1 respectively so they disappear from the equation.
p. 47, Table 2.3, line 7	`chisq.pdf(x, ν)`	`chisq.dist(x, ν, false)`
p. 48, line 3	or *yes* or *no*	or *yes* **and** *no*
p. 48, line 5	It is assumed that there are *k* trials…	It is assumed that there are ***n*** trials…
p. 50, line 9-12	It should be noted that, as for the binomial distribution from which it is derived, it is possible to approximate the Poisson distribution using the standard normal distribution if $\lambda > 5$. In this case,	It should be noted that, as for the binomial distribution from which it is derived, it is possible to approximate the Poisson distribution using the standard normal **distribution. If $\lambda > 5$,**

---

[2] dy should be replaced by dz.

Page, line	Current Form	Correction
p. 51, bullet #3	mean square error	mean squared error
p. 52, Equation (2.32)	$\text{MSE}\left(\hat{\theta}\right)=\text{E}\left(\left(\hat{\theta}-E\left(\hat{\theta}\right)\right)^{2}\right)=\sigma_{\hat{\theta}}^{2}+\delta^{2}$	$\text{MSE}\left(\hat{\theta}\right)=\text{E}\left(\left(\hat{\theta}-\theta\right)^{2}\right)=\sigma_{\hat{\theta}}^{2}+\delta^{2}$
p. 55, Equation (2.41)	$\ell\left(\theta\middle\|\vec{x}\right)=\log L\left(\theta\middle\|\vec{x}\right)=\sum_{i=1}^{n} f\left(x_i,\theta\right)$	$\ell\left(\theta\middle\|\vec{x}\right)=\log L\left(\theta\middle\|\vec{x}\right)=\sum_{i=1}^{n}\log f\left(x_i,\theta\right)$ The **log** is missing from the summation.
p. 63, bullet #3, line 2	…for a small initial confidence intervals.	*remove these words.*
p. 64, line 12	…variance is known (in which it should be used in lieu of $\hat{\sigma}$) or $n > 30$.	…variance is known (in which **case** it should be used in lieu of $\hat{\sigma}$) or $n > 30$.
p. 66, Example 2.10, Equation	$t_{computed}=\frac{1.56-1}{\frac{0.77}{\sqrt{7}}}=1.92$	$t_{computed}=\frac{1.56-1}{\frac{0.7\underline{8}}{\sqrt{7}}}=1.9\underline{0}$
p. 68, Example 2.13, line 5	…during 40 h of operation.	…during **50** h of operation.

Page, line	Current Form	Correction
p. 69, §2.7.6.1, line 2	It all cases,…	**In** all cases,…
p.73, lines 1/2	Example 2.15: Testing the Difference in Means—Unknown, Common Mean	Example 2.15: Testing the Difference in Means—Unknown, Common **Variance**
p. 77 (last line of the page)	(not centred)	(centred; The formula should be centred on the page.)
p. 80, bullet #13	mean square error	mean squared error
p. 81, Equation (2.68)	$x \geq 0$	$\boldsymbol{y} \geq 0$
p. 82, bullet #27	mean square error	mean squared error
p. 95 ff, Theorem 3.2 (bis)	Theorem 3.2: *Under the assumption*	Theorem 3.3: *Under the assumption* (and this implies that the numbering is one off for the rest of the theorems.)
p. 95, above Eq. (3.13)	The variance of the parameters can be written as	From Equation (2.20), the variance of the parameters can be written as (*clarifying how the variance is obtained*)

Page, line	Current Form	Correction
p. 96, Equation (3.18)	$\sigma^2_{\hat{\beta}} = E\left(\varepsilon\varepsilon^T\right) = \sigma^2 \mathcal{I}$	$\underline{\sigma^2_{\varepsilon}} = E\left(\varepsilon\varepsilon^T\right) = \sigma^2 \mathcal{I}$
p. 96, Equation (3.20)	$\hat{\beta}_i \pm t_{n-m,1-\frac{\alpha}{2}} \hat{\sigma} \sqrt{\left(\mathcal{A}^T \mathcal{A}\right)^{-1}_{ii}}$	$\hat{\beta}_i \pm t_{1-\frac{\alpha}{2},m-n} \hat{\sigma} \sqrt{\left(\mathcal{A}^T \mathcal{A}\right)^{-1}_{ii}}$ (for consistency and correcting $n - m$ to $m - n$)
p. 101, line 15	where α is the alpha error	where α is the **α-**error
p. 101, line 16	using Eq. (138)	using Eq. **(3.44)**
p. 106, line 4	$t_{0.975,\ 7-1} = 2.967$	$t_{0.975,\ 7-1} = 2.447$
p. 106, Equation (3.66)	$28.3529 \pm 2.9687\left(0.310\ 43\right)\sqrt{0.714\ 28}$ $28.3529 \pm 0.7789 \frac{\text{kg}}{\text{min} \bullet \text{m}^{0.5}}$	$28.3529 \pm 2.447\left(0.310\ 43\right)\sqrt{0.714\ 28}$ $28.3529 \pm 0.642 \frac{\text{kg}}{\text{min} \bullet \text{m}^{0.5}}$
p. 106, line 10	0.8	0.6
p. 106, line 11	28.4±0.8 $\text{kg} \cdot \text{min}^{-1} \cdot \text{m}^{-0.5}$	28.4±0.6 $\text{kg} \cdot \text{min}^{-1} \cdot \text{m}^{-0.5}$

Page, line	Current Form	Correction
p. 106, Equation (3.67)	$\vec{x}_0 \hat{\vec{\beta}} \pm t_{1-\frac{\alpha}{2},m-n}\,\hat{\sigma}\sqrt{\vec{x}_0(\mathcal{A}^T\mathcal{A})^{-1}\vec{x}_0^T}$ $\sqrt{0.225}\left(28.3529\right)\pm$ $2.9687\left(0.310\ 43\right)\sqrt{\sqrt{0.225}\left(0.714\ 28\right)\sqrt{0.225}}$ $13.448\ 96\pm 0.3694\frac{\text{kg}}{\text{min}}$	$\vec{a}_{\vec{x}_d}\hat{\vec{\beta}} \pm t_{1-\frac{\alpha}{2},m-n}\,\hat{\sigma}\sqrt{\vec{a}_{\vec{x}_d}(\mathcal{A}^T\mathcal{A})^{-1}\vec{a}_{\vec{x}_d}^T}$ $\sqrt{0.225}\left(28.3529\right)\pm$ $2.447\left(0.310\ 43\right)\sqrt{\sqrt{0.225}\left(0.714\ 28\right)\sqrt{0.225}}$ $13.448\ 96\pm 0.304\frac{\text{kg}}{\text{min}}$
p. 106, line 14	13.5±0.4 kg·m^{-1}	13.**4**±0.**3** kg·m^{-1}
p. 106, Equation (3.68)	$\vec{x}_0 \hat{\vec{\beta}} \pm t_{1-\frac{\alpha}{2},m-n}\,\hat{\sigma}\sqrt{1+\vec{x}_0(\mathcal{A}^T\mathcal{A})^{-1}\vec{x}_0^T}$ $\sqrt{0.225}\left(28.3529\right)\pm$ $2.9687\left(0.310\ 43\right)\sqrt{1+\sqrt{0.225}\left(0.714\ 28\right)\sqrt{0.225}}$ $13.448\ 96\pm 0.9929\frac{\text{kg}}{\text{min}}$	$\vec{a}_{\vec{x}_d}\hat{\vec{\beta}} \pm t_{1-\frac{\alpha}{2},m-n}\,\hat{\sigma}\sqrt{1+\vec{a}_{\vec{x}_d}(\mathcal{A}^T\mathcal{A})^{-1}\vec{a}_{\vec{x}_d}^T}$ $\sqrt{0.225}\left(28.3529\right)\pm$ $2.447\left(0.310\ 43\right)\sqrt{1+\sqrt{0.225}\left(0.714\ 28\right)\sqrt{0.225}}$ $13.448\ 96\pm 0.8184\frac{\text{kg}}{\text{min}}$
p. 106, line 18	14±1 kg·m^{-1}	1**3**±**0.8** kg·m^{-1}
p. 106, Equation (3.69)	$SSE=\hat{\sigma}^2\left(m-n\right)=\left(0.31041\right)^2\left(7-1\right)=0.5781$	$SSE=\hat{\sigma}^2\left(m-n\right)=\left(0.310\ 41\right)^2\left(7-1\right)=0.5781$ Proper formatting of the variance.
Table 3.2 (fifth cell)	Prounanced Tails	**Pronounced** Tails

Page, line	Current Form	Correction
p. 111, 17	the letters correspond to the graphs shown in Table 3.4 as problem graphs	**the letters correspond to the problem graphs shown in Table 3.4**
Table 3.4	plot(s)	**graph(s)**
p. 114, line 28	$\hat{R} = 28.4 \pm 0.8 \frac{\text{kg}}{\text{min}\bullet\text{m}^{0.5}}$	$\hat{R} = 28.4 \pm 0.6 \frac{\text{kg}}{\text{min}\bullet\text{m}^{0.5}}$
p. 119, Example 3.2	…(consider replicates 2 and 3 of run 2).	…(consider replicates 2 and 3 of **Run 3**).
p. 120, §3.4, line 15	Levenberg–Marquardt algorithm optimisation algorithms	Levenberg–Marquardt algorithms
p. 122, line 30	The standard deviation, $\hat{\sigma}$,	The standard deviation **for this model**, $\hat{\sigma}$,
p. 124, 2nd line after Equation (3.108)	ln$A$ = 5.700 84±0.007 62 and $-E_aR^{-1}$ = −178.84±2.208 78	ln$A$ = 5.700 84±0.006 53 and $-E_aR^{-1}$ = −178.84±1.893 51

Page, line	Current Form	Correction
p. 124, 2nd paragraph after Equation (3.108)	This gives $\hat{A}_{lower} = e^{5.700\ 84 - 0.007\ 62} = 296.8$ and $\hat{A}_{lower} = e^{5.700\ 84 + 0.007\ 62} = 301.41$.	This gives $\hat{A}_{lower} = e^{5.700\ 84 - 0.006\ 53} = 297.2$ and $\hat{A}_{upper} = e^{5.700\ 84 + 0.006\ 53} = 301.08$.
p. 124, 3rd paragraph after Equation (3.108)	The confidence interval becomes 2.208 78×−8.314 = (−)18.36. Therefore, the confidence interval for E_a is 1,490±18 J·mol^{-1}.	The confidence interval becomes 1.893 51×−8.314 = (−)15.74. Therefore, the confidence interval for E_a is 1,490±16 J·mol^{-1}.
p. 125, 2nd line	$\hat{E}_a$ = 1,490±21 J·mol^{-1}	$\hat{E}_a$ = 1,490±18 J·mol^{-1}
p. 125, Figure 3.4 (bottom, left)		*see Appendix I: Correction to Figure 3.4 (bottom, left) for the correct figure.*

Page, line	Current Form	Correction
p. 125, starting from the second line after Equation (3.110)	Secondly, Fig. 3.4 show normal probability plots…This shows the importance of selecting an appropriate method for the given problem.	*see Appendix I: Correction to Figure 3.4 (bottom, left) for the corrected text.*
p. 131, 3	The confidence interval for T = 105°F is more reliable	The **prediction** for T = 105°F is more reliable
p. 135, one line above Equation (3.113)	the properties of the system	the water level in the tank
p. 135, one line below Equation (3.113)	height	water level
p. 136, Table 3.13, caption	tank height	water level

Page, line	Current Form	Correction
p. 136, Table 3.13, second column header	Height	level
p. 138, Equation (3.A7)	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2}\hat{\sigma}\sqrt{\frac{1}{m}+\frac{\left(x_d-\frac{1}{m}\sum x_i\right)}{s_x^2}}$	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2}\hat{\sigma}\sqrt{\frac{1}{m}+\frac{\left(x_d-\frac{1}{m}\sum x\right)^2}{s_x^2}}$
p. 138, Equation (3.A8)	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2}\hat{\sigma}\sqrt{1+\frac{1}{m}+\frac{\left(x_d-\frac{1}{m}\sum x_i\right)}{s_x^2}}$	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2}\hat{\sigma}\sqrt{1+\frac{1}{m}+\frac{\left(x_d-\frac{1}{m}\sum x\right)^2}{s_x^2}}$
p. 139, Section A.2	The ordinary, least-squares problem can be solved by first computing the following two quantities	The **weighted**, least-squares problem can be solved by first computing the following two quantities
p. 139, Equation (3.A18)	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2}\hat{\sigma}_w\sqrt{\frac{1}{\left(\sum w_i\right)}+\frac{\left(x_d-\frac{1}{\left(\sum w_i\right)}\sum w_i x_i\right)}{s_{x_w}^2}}$	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2}\hat{\sigma}_w\sqrt{\frac{1}{\left(\sum w\right)}+\frac{\left(x_d-\frac{1}{\left(\sum w\right)}\sum wx\right)^2}{s_{x_w}^2}}$
p. 139, Equation (3.A19)	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2-n'}\hat{\sigma}_w\sqrt{\frac{1}{w_d}+\frac{1}{\left(\sum w_i\right)}+\frac{\left(x_d-\frac{1}{\left(\sum w_i\right)}\sum w_i x_i\right)}{s_x^2}}$	$\hat{y} \pm t_{1-\frac{\alpha}{2},m-2-n'}\hat{\sigma}_w\sqrt{\frac{1}{w_d}+\frac{1}{\left(\sum w\right)}+\frac{\left(x_d-\frac{1}{\left(\sum w\right)}\sum wx\right)^2}{s_x^2}}$

Page, line	Current Form	Correction
p. 142, Equation (4.1)	$S_{\beta_i} = \dfrac{\partial \vec{f}\left(\vec{x}, \vec{\beta}\right)}{\partial \beta_i}$	$\vec{S}_{\beta_i} = \dfrac{\partial \vec{f}\left(\vec{x}, \vec{\beta}\right)}{\partial \beta_i}$ (for consistency)
p. 148, Equation (4.5)	$y_i = \beta_0 + \sum_{j=1}^{k} \sum_{d=1}^{l-1} \beta_{j^d} x_j^d + \prod_{\substack{\text{in twos,}\\ \text{threes,}\\ \ldots,\\ \text{groups of } l}} \left(\beta_{..} \sum_{j=1}^{k} \sum_{d=1}^{l-1} x_j^d \right) + e_i$	$y_i = \beta_0 + \sum_{j=1}^{k} \sum_{d=1}^{l-1} \beta_{j^d} x_j^d + \prod_{\substack{\text{in twos,}\\ \text{threes,}\\ \ldots,\\ \text{groups of } k}} \left(\beta_{..} \sum_{j=1}^{k} \sum_{d=1}^{l-1} x_j^d \right) + e_i$
p. 148, line 4	until a single group of all *l* parameters	until a single group of all $\underline{\boldsymbol{k}}$ parameters
p. 151, line above Equation (4.18)	Eq. (3.17)	Eq. ($\underline{\mathbf{4}}$.17)

Page, line	Current Form	Correction
p. 156, first two lines	see footnote [3]	see footnote [4] The equations are too large to show otherwise and all entries are wrong.
p. 156f, Example 4.2, b), c), and d)	*Some of the comments for question b), c), and d) are not correct given the error in the computation of the parameter values.*	*Please see Appendix I: Correction for Example 4.2 for the updated version.*
p. 159, §4.5.3.1, first sentence of paragraph 3	where x is the divisor and y is the dividend (or base)	where x is the **dividend** and y is the **divisor** (or base)

---

[3]
$$\hat{\vec{\beta}} = 2^{-4}\mathcal{A}^T\vec{y}$$
$$= \begin{bmatrix} 101 & 5.19 & -0.813 & -2.19 & 3.06 & -0.0625 & -7.69 & -0.438 & 0.813 & 0.813 & \dots \\ & & & \dots & & -0.313 & 0.313 & -0.188 & -0.0625 & 0.188 & -0.313 \end{bmatrix}^T \text{ (original)}$$

[4]
$$\hat{\vec{\beta}} = 2^{-4}\mathcal{A}^T\vec{y}$$
$$= \begin{bmatrix} 70.06 & 10.81 & 1.56 & 4.94 & 7.31 & 0.0625 & -9.06 & 8.31 & 1.19 & -0.188 & \dots \\ & & & \dots & & -0.563 & 0.938 & 2.063 & -0.813 & -1.32 & 0.688 \end{bmatrix}^T \text{ (corrected)}$$

Page, line	Current Form	Correction
p. 161f, §4.5.3.4	n	k (replace all n in this section by k, the number of factors)
p. 162, line 28	Since n is odd…	Since $k = 5$ and is odd… (to clarify the situation!)
p. 171, lines 5/6	Using the F-test approach to determining the significant parameters, find a reduced model and analyse its residuals.	Using the F-test approach to determining the significant parameters, find a reduced model and analyse its residuals **<u>(α = 0.05)</u>**.
p. 173, Table 4.4	F-critical, model F-test	F-critical, model $F_{0.95, 15, 16}$ F-test, model F-critical, parameter $F_{0.95, 1, 16}$ 4.49 (*added*)
p. 174, Example 4.9, line 4	63%	67% (It should be $100 \times R^2$.)
p. 174, Example 4.9, Table 4.5		F_i for β_{15} and β_{35} should be bolded, as they are significant.
p. 174, Example 4.9, Table 4.5	F-critical, model F-test	F-critical, model $F_{0.95, 1, 30}$ F-test, model

Page, line	Current Form	Correction
p. 175, Example 4.9, Figure 4.6 (bottom)		*see Appendix III: Correction for Example 4.9, Figure 4.6 (bottom).* (The figure has been updated to have properly labelled axes. There are no overall changes in the figures.)
p. 175, Example 4.9, Figure 4.7		*see Appendix IV: Correction for Example 4.9, Figure 4.7.* (The figures have been updated to make clear that the correct reduced model was used. There are no overall changes in the figures.)
p. 175, line 4	is now 0.91	is now <u>0.88</u>
p. 180, Equation (4.39)	$\gamma_{ji} = \sum_{k=0}^{\frac{j}{2}} \beta_{jk} x_i^k$	$\gamma_{ji} = \sum_{k=0}^{\frac{j}{2}} \beta_{jk} x_i^{2k}$ The superscript on the *x* should be ***2k*** rather than *k*.
p. 182, line 6 (right after Eq. 4.51)	$\gamma_{11} + \gamma_{12} + \gamma_{13} = 0$	$\gamma_{\underline{\mathbf{2}}1} + \gamma_{\underline{\mathbf{2}}2} + \gamma_{\underline{\mathbf{2}}3} = 0$

Page, line	Current Form	Correction
p. 182, line 7 (right after Eq. 4.51)	(6 unknowns, but 5 equations)	(**5** unknowns, but **4** equations)
p. 182, line 8 (right before Eq. 4.52)	$\gamma_{13} = \gamma_{11} = 1$	$\gamma_{\underline{\mathbf{2}}3} = \gamma_{\underline{\mathbf{2}}1} = 1$
p. 184, Equation (4.55), 4th line	$\gamma_{13} = \beta_{11}x_4 = \beta_{11}(1)$	$\gamma_{14} = \beta_{11}x_4 = \beta_{11}(1)$ It should be $\gamma_{\mathbf{14}}$ rather than $\gamma_{\mathbf{13}}$.
p. 184, Equation (4.56)	$\gamma_{11} = -1, \gamma_{12} = -{}^{1}\!/_{3}, \gamma_{12} = {}^{1}\!/_{3}, \gamma_{13} = 1$ $\beta_{11} = 1$	$\gamma_{11} = -1, \gamma_{12} = -{}^{1}\!/_{3}, \gamma_{13} = {}^{1}\!/_{3}, \gamma_{14} = 1$ $\beta_{11} = 1$ The second γ_{12} should read as $\gamma_{\mathbf{13}}$ and the γ_{13} as $\gamma_{\mathbf{14}}$**.**
p. 185, line 7 (right after Eq. 4.61)	$\gamma_{11} + \gamma_{12} + \gamma_{13} + \gamma_{14} = 0$	$\gamma_{\underline{\mathbf{2}}1} + \gamma_{\underline{\mathbf{2}}2} + \gamma_{\underline{\mathbf{2}}3} + \gamma_{\underline{\mathbf{2}}4} = 0$

Page, line	Current Form	Correction
p. 185, line 8 (right after Eq. 4.61)	$\gamma_{13} = \gamma_{12}$ and $\gamma_{14} = \gamma_{11}$	$\gamma_{\underline{\mathbf{2}}3} = \gamma_{\underline{\mathbf{2}}2}$ and $\gamma_{\underline{\mathbf{2}}4} = \gamma_{\underline{\mathbf{2}}1}$
p. 185, line 9 (right before Eq. 4.62)	$\gamma_{13} = \gamma_{12} = -1$	$\gamma_{\underline{\mathbf{2}}3} = \gamma_{\underline{\mathbf{2}}1} = -1$
p. 185, line 14 (right after Eq. 4.63)	$\gamma_{11} + \gamma_{12} + \gamma_{13} + \gamma_{14} = 0$	$\gamma_{\underline{\mathbf{3}}1} + \gamma_{\underline{\mathbf{3}}2} + \gamma_{\underline{\mathbf{3}}3} + \gamma_{\underline{\mathbf{3}}4} = 0$
p. 185, line 16 (right after Eq. 4.63)	$\gamma_{13} = -\gamma_{12}$ and $\gamma_{14} = -\gamma_{11}$	$\gamma_{\underline{\mathbf{3}}3} = -\gamma_{\underline{\mathbf{3}}2}$ and $\gamma_{\underline{\mathbf{3}}4} = -\gamma_{\underline{\mathbf{3}}1}$
p. 185, line 18 (right after Eq. 4.63)	$\gamma_{11} = -1, \ldots, \gamma_{14} = 1$, and $\gamma_{12} = 1$	$\gamma_{\underline{\mathbf{3}}1} = -1, \ldots, \gamma_{\underline{\mathbf{3}}4} = 1$, and $\gamma_{\underline{\mathbf{3}}2} = 1$

Page, line	Current Form	Correction
p. 185, line 19 (right before Eq. 4.64)	$\gamma_{13} = -1$	$\gamma_{\underline{\mathbf{3}}3} = -1$
p. 188, Equation (4.70)	$\tilde{x}_1 = \frac{x_1 - 0.5(14+10)}{0.5(14-10)} = 0.5x_1 - 6$ $\tilde{x}_2 = \frac{x_3 - 0.5(25+30)}{0.5(30-25)} = 0.4x_2 - 11$ $\tilde{x}_3 = \frac{x_2 - 0.5(200+250)}{0.5(250-200)} = 0.04x_3 - 9$	$\tilde{x}_1 = \frac{x_1 - 0.5(14+10)}{0.5(14-10)} = 0.5x_1 - 6$ $\tilde{x}_2 = \frac{x_2 - 0.5(25+30)}{0.5(30-25)} = 0.4x_2 - 11$ $\tilde{x}_3 = \frac{x_3 - 0.5(200+250)}{0.5(250-200)} = 0.04x_3 - 9$ In the second and third equations, after the first equals sign, x_3 (in the second equation) should be $\boldsymbol{x_2}$ and x_2 (in the third equation) should be $\boldsymbol{x_3}$.
p. 189, Section 4.7.4.4	$L_2(\tilde{x}_1) = 3\tilde{x}_1 - 2$	$L_2(\tilde{x}_1) = 3\tilde{x}_1^{\mathbf{2}} - 2$

Page, line	Current Form	Correction
p. 189, Equation (4.72)[5]	$\mathcal{A}_{part} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -2 \\ 1 & 0 & -1 & -1 & 0 & 0 & 1 & 0 & -2 & 2 & 2 & 2 \\ 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 & -2 & 2 & -2 & 2 \\ 1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & -2 & -2 & 2 & -2 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & -2 & -2 & -2 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	$\mathcal{A}_{part} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 0 & -1 & -1 & 0 & 0 & 1 & 0 & -2 & 2 & 2 & -2 \\ 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 & -2 & 2 & -2 & 2 \\ 1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & -2 & -2 & 2 & 2 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & -2 & -2 & -2 & -2 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
p. 189, Equation (4.74)[6]	$\vec{y} = [\cdots \quad 1 \quad 1 \quad 3 \quad 1 \quad 5 \quad 4 \quad 9 \quad 6 \quad 11]^T$	$\vec{y} = [\cdots \quad 1 \quad 1 \quad \mathbf{1} \quad \mathbf{3} \quad 5 \quad 4 \quad \mathbf{6} \quad \mathbf{9} \quad 11]^T$
p. 191, Figure 4.8	β_{13}	$\beta_{1\underline{\mathbf{2}}}$

---

[5] The last column of the matrix was shifted up by one row. The corrected matrix is provided. All subsequent calculations are still correct.

[6] Only the last 9 elements of this vector are shown. Two groups of two digits were transposed. All subsequent calculations are still correct.

Page, line	Current Form	Correction
p. 191, Figures 4.10, 4.11	*There seems to be some minor differences in the graphs due to ordering of variables. The interpretation and final results are not affected.*	*See Appendix V: Updating Figures 44 and 45 for the Example in §4.7.4*
p. 196, Equation (4.90)	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_{12} + \beta_c x^2$	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_c x^2$ The third term, $\beta_{12}x_{12}$, should be $\boldsymbol{\beta_{12}x_1x_2}$.
p. 198, line 17	F-statistic is 15.9 (> 6.4	F-statistic is 47.8 (> 5.14
Figures 4.13, 4.14, 4.15	*clarifying the axis labels*	*See Appendix IV: Updating Figures 4.13, 4.14, and 4.15 for the Example in §4.8.4*
p. 201, §4.9.2, bullet 4	largest value of $N\mathrm{var}(\hat{y})/\sigma^2$	largest value of $N\mathrm{var}(\hat{y})/\sigma^2$**, where N is the number of design points in the experiment, for example, for a full factorial design, $N = l^k$.** *(Clarifying the meaning of the variable.)*
p. 213, line 6	$E(x_t)$	**E(**x_t**)** (*Only the value inside the bracket should be in italics*.)
p. 216, line 8	There are two types of differencing: true and periodic.	There are two types of differencing: **true** and **periodic**. (*Adding bolding to emphasis the points*.)

Page, line	Current Form	Correction
p. 220, Equation (5.27)	$A_Q\left(z^{-s}\right)=1+\sum_{i=1}^{P}\alpha_{si}z^{-si},$ $B_P\left(z^{-s}\right)=1+\sum_{i=1}^{Q}\beta_{si}z^{-si}.$	$\underline{A_P}\left(z^{-s}\right)=1+\sum_{i=1}^{P}\alpha_{si}z^{-si},$ $\underline{B_Q}\left(z^{-s}\right)=1+\sum_{i=1}^{Q}\beta_{si}z^{-si}.$
p. 221, right after Equation (5.27)	The pure seasonal autoregressive model…	The pure seasonal **moving average** model…
p. 226, computation of σ_y, 3rd line	$=\sigma^2+0.5^2\sigma^2+0.1\sigma^2=$	$=\sigma^2+0.5^2\sigma^2+\underline{0.1^2}\sigma^2=$
p. 252, line 4	…model (Fig. 5.14).	…model. Fig. 5.14 shows the normal probability plot and the autocorrelation plot of the residuals. (*Adding clarifying text.*)
p. 254, Equation (5.102)	$\sigma^2_{t+1\|t}=\gamma\left(0\right)-\vec{\gamma}_t^T\Gamma_t^{-1}\vec{\gamma}_t$	$\sigma^2_{t+\underline{\tau}\|t}=\gamma\left(0\right)-\vec{\gamma}_t^T\Gamma_t^{-1}\vec{\gamma}_t$ (The subscript on the σ should be $t+\underline{\tau}\mid t$ rather than $t+\underline{1}\mid t$.)
p. 255, Equation (5.105)	$y_{t+\tau}=E\left(y_{t+\tau}\mid y_t,y_{t-1},y_{t-2},\ldots\right)=\sum_{j=\tau}^{\infty}h_je_{t+\tau-j}$	$\underline{\hat{y}}_{t+\tau\|\underline{t}}=E\left(y_{t+\tau}\mid y_t,y_{t-1},y_{t-2},\ldots\right)=\sum_{j=\tau}^{\infty}h_je_{t+\tau-j}$

Page, line	Current Form	Correction
p. 256, 1-step-ahead	$\hat{y}_{101\|100} = 0.436e_{99} - 0.293e_{98} - 0.763e_{97}$ $= 0.436(0.5044) - 0.293(-0.4402) - 0.763(-1.3221)$ $= 1.3577$	$\hat{y}_{101\|100} = 0.436e_{100} - 0.293e_{99} - 0.763e_{98}$ $= 0.436(0.8410) - 0.293(0.5044) - 0.763(-0.4402)$ $= 0.5548$
p. 256, 1-step-ahead	…gives a 95% confidence interval as 1.4±2.1.	…gives a 95% confidence interval as 0.6±2.1.
p. 256, 2-step-ahead	$\hat{y}_{102\|100} = -0.293e_{99} - 0.763e_{98}$ $= -0.293(0.5044) - 0.763(-0.4402)$ $= 0.1881$	$\hat{y}_{102\|100} = -0.293e_{100} - 0.763e_{99}$ $= -0.293(0.8410) - 0.763(0.5044)$ $= -0.6312$
p. 256, 2-step-ahead	$\hat{\sigma}^2_{102\|100} = 1.0807\left(1^2 + 0.436^2\right) = 1.2708$	$\hat{\sigma}^2_{102\|100} = 1.0807\left(1^2 + 0.436^2\right) = \underline{1.286}$
p. 256, 2-step-ahead	95% confidence interval as 0.2±2.5	95% confidence interval as **−0.6**±2.5
p. 268, Equation (5.132)	$\Sigma_{t-1\|t} = \mathcal{A}\Sigma_{t-1\|t}\mathcal{A}^T + \Sigma_\omega$	$\Sigma_{t\|t-1} = \mathcal{A}\Sigma_{t-1\|t-1}\mathcal{A}^T + \Sigma_\omega$
p. 268, Equation (5.134)	$\mathcal{K}_t = \Sigma_{t-1\|t}\mathcal{C}^T\left(\mathcal{C}\Sigma_{t-1\|t}\mathcal{C}^T + \Sigma_e\right)$	$\mathcal{K}_t = \Sigma_{t\|t-1}\mathcal{C}^T\left(\mathcal{C}\Sigma_{t\|t-1}\mathcal{C}^T + \Sigma_e\right)^{-1}$
p. 269, lines 10, 16	the *z*-value of	the ***Z***-value of

Page, line	Current Form	Correction
p. 274, 3.c, line 3	*Elecotroacustics*	*Electroacoustics*
pp. 287 – 289	*m*-step	$\underline{\boldsymbol{\tau}}$-step
p. 288, line 4	where some L is a rational function of z^{-1}	where **L is some rational** function of z^{-1}
p. 293, Equation (6.31) p. 293, line after Equation (6.31)	$t \mid t-1$	$\underline{\boldsymbol{t+1 \mid t}}$ (Note both formulations are equivalent, but for consistency, we should use the corrected form.)
p. 294, Equation (6.41)	$\operatorname{var}\left(\varepsilon_t\left(\vec{\theta},\hat{\vec{\theta}}\right)\right)=\operatorname{var}\left(\Phi_e\left(z^{-1},\vec{\theta},\hat{\vec{\theta}}\right)u_t\right)$ $+\operatorname{var}\left(\left(\Phi_e\left(z^{-1},\vec{\theta},\hat{\vec{\theta}}\right)-\mathcal{I}\right)e_t\right)$ $+\operatorname{var}\left(e_t\right)$	$\operatorname{var}\left(\varepsilon_t\left(\vec{\theta},\hat{\vec{\theta}}\right)\right)=\operatorname{var}\left(\boldsymbol{\Phi_u}\left(z^{-1},\vec{\theta},\hat{\vec{\theta}}\right)u_t\right)$ $+\operatorname{var}\left(\left(\Phi_e\left(z^{-1},\vec{\theta},\hat{\vec{\theta}}\right)-\mathcal{I}\right)e_t\right)$ $+\operatorname{var}\left(e_t\right)$ The first $\mathbf{\Phi}_{\boldsymbol{e}}$ should be $\mathbf{\Phi}_{\boldsymbol{u}}$ instead.

Page, line	Current Form	Correction
p. 303, first paragraph of §6.4	In such cases, using and understanding routine operating data **are** important.	In such cases, using and understanding routine operating data **is** important.
p. 305, line 12	direction identification	**direct identification**
p. 310, §6.5.1, second paragraph	Weiner-Hammerstein models are useful with the actuators or sensors have…	**Wiener**-Hammerstein models are useful with the actuators or sensors have…
p. 311, Fig. 6.2		(See Figure 1 for the corrected image, which clarifies the relationship between the flows in the 4 tanks.)
p. 314, Figure 6.10	(cm/s)	(cm^3/s)
p. 316, Equation (6.82), second line	$\frac{\left(7.8\times10^{-4}\pm2\times10^{-5}\right)z^{-2}}{1-\left(1.664\pm0.007\right)z^{-1}+\left(0.695\pm0.007\right)z^{-1}}u_2$	$\frac{\left(7.8\times10^{-4}\pm2\times10^{-5}\right)z^{-2}}{1-\left(1.664\pm0.007\right)z^{-1}+\left(0.695\pm0.007\right)\underline{z^{-2}}}u_2$
p. 324, line 14	This section gives	This **appendix** gives
p. 340, Table 7.3	start, *	**star,** *

Page, line	Current Form	Correction
p. 341, Table 7.4	cumulative density function	probability density function (*This occurs a few times in the table.*)
p. 341, Table 7.4, row 5	`normcdf(x,m,s) |` Calculates the normal cumulative density function…	`normpdf(x,m,s)|` Calculates the **probability density function for the normal distribution**…
p. 343, Table 7.7, first line	`=regress(b`	`=regress(y`
p. 347, Table 7.10, second line, first column	`mARab=ar(z,[na,nb])`	`mARa=ar(z,[na])`
p. 347, Table 7.10, second line, second column	orders `na` and `nb`… an `idpoly` object, `mARab`	order `na`… an `idpoly` object, `mARa`
p. 353, line 24	`%Obtain the autocorrelation values`	`%Obtain the crosscorrelation values`

Page, line	Current Form	Correction
p. 356, third computer text box	`A=A([1, size(A,1)],:);`	`A=A([1, 3:size(A,1)],:);` (**3:**)
p. 359, first line of second computer text box	`%Script for solving linear regression problems in MATLAB`	`%Script for solving nonlinear regression problems in MATLAB` (**nonlinear**)
p. 360, Figure 7.2, caption	Fig. 7.2 Linear regression example:	Fig. 7.2 **Nonlinear** regression example:
p. 361, computer text box	`cm/s`	`cm^3/s` (**^3**)
p. 363	Microsoft Office 2013	Microsoft Office **2016**
p. 374, §8.5	In order to start Solver, in Excel 2007 or newer,… Solver should be there as shown in Fig. 8.7.	In order to start **the Data Analysis Add-In**, in Excel 2007 or newer,… **The Data Analysis Add-In** should be there as shown in Fig. 8.7.
p. 379, item 3)	Median – Q1, Q3, and Maximum – Q3	Median – Q1, **Q3 − Median**, and Maximum – Q3

Page, line	Current Form	Correction
p. 379, item 4)	Q3	**<u>Q3 – Median</u>**
p. 399, Chapter 2, 15)	15) T;	15) <u>F</u>;
p. 401, last line	24) left graph: 10	24) left graph: **<u>9</u>**
p. 405, line 29	*Elecotroacustics*	*Electroacoustics*

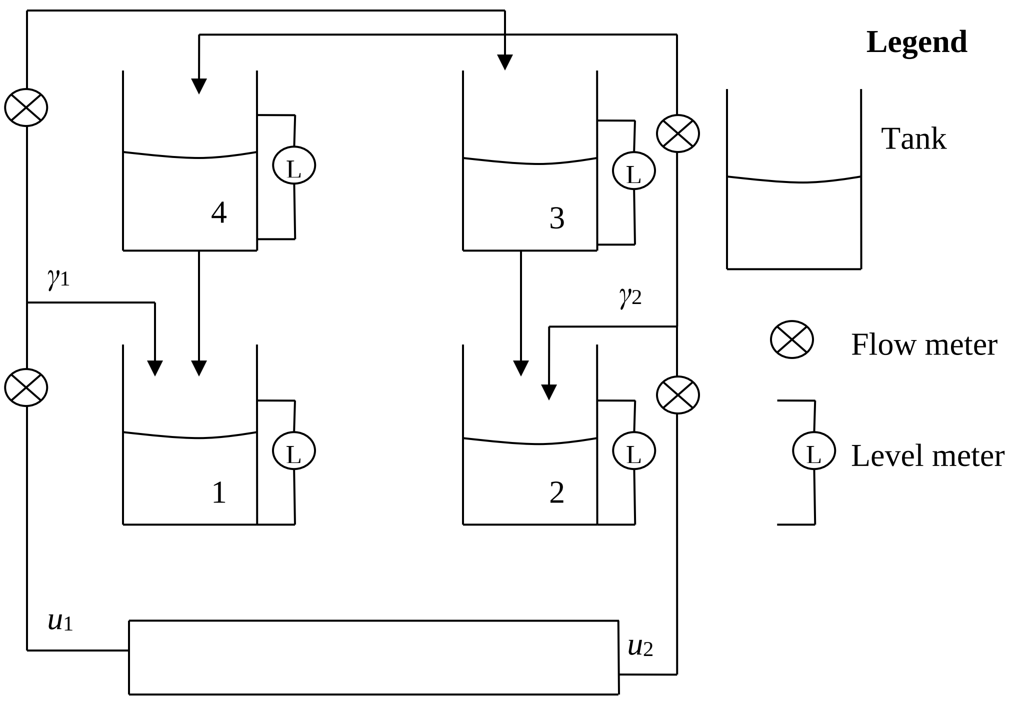
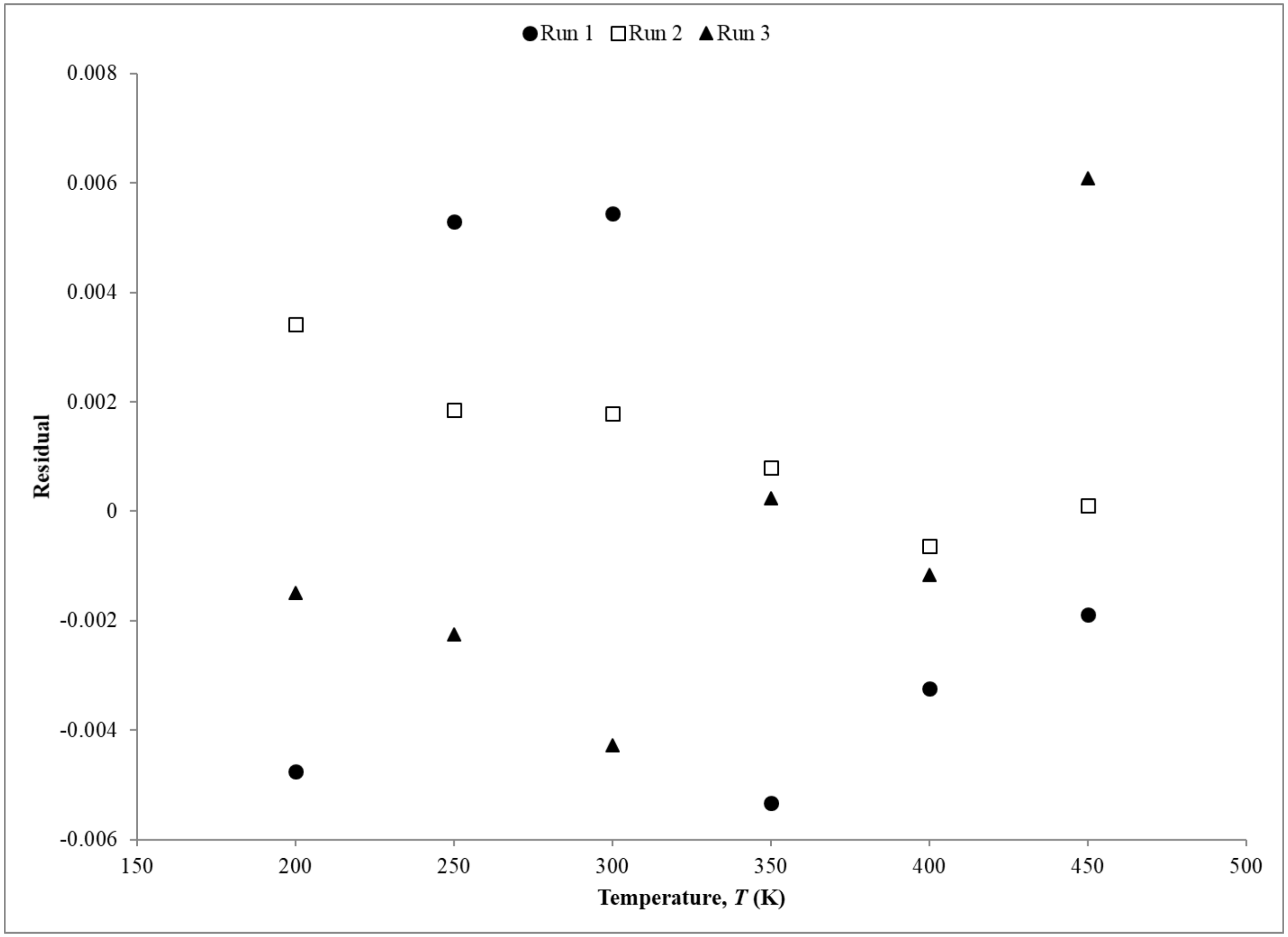


Figure 1: Corrected Fig. 6.2

Yuri A.W. Shardt would like to thank his students in the *Intelligent Regelung* (en: *Smart Control*) course, Thomas Donnelly, Alexandru Vasile, and Heiko Weiß for pointing out typos and unclear sections.

Appendix I: Correction to Figure 3.4 (bottom, left) and corrected text



Corrected Text:

Secondly, Fig. 3.4 shows **the** normal probability plots and the residuals as a function of the temperature for both cases. From the normal probability plots, it would seem that the residuals for both models are quite similar. On the other hand, there do seem to be more abnormal points in the linearised model case, suggesting that the residuals may violate the assumption of normality. Examining the residual as a function of temperature plots shows some interesting results. Firstly, for the linearised model, Run **2 has a different behaviour from the other two runs.** Secondly, there are few, if any, **positive** deviations compared with the large number of **negative** deviations. Based solely on the linearised model, one would have to conclude that Run **2** was abnormal, and the collection of the data would warrant additional scrutiny. On the other hand, when the nonlinear case is used, the results are quite different and a different pattern emerges. First, Run **2** is no longer usual, and there are now both positive and negative residuals in equal magnitude. Second, it would seem that the residuals depend on the temperature with a lower value around 350 to 400 K and higher values at the extremes. Since Arrhenius's equation is an accepted model for the observed behaviour, this feature could potentially be attributed to issues in experimental design, that is, the conditions and methods by which the data were obtained, for example, faulty measurements or an incomplete procedure.

It is interesting to note that, although both the linear and nonlinear methods provided similar parameter estimates and confidence intervals, the residual analysis is quite different. In the linear case, it would be concluded that Run **2** had some abnormal residuals and would require additional analysis. In the nonlinear case, it would be concluded that there seems to be some temperature dependency of the residuals. This shows the importance of selecting an appropriate method for the given problem.

Appendix II: Correction for Example 4.2, Parts b), c), and d)

b) A normal probability plot of the effects is shown in Figure 37. **The effects that lie far from the expected normal distribution values are those that are significant because they are not chance values. The most significant effects have been circled and labelled. Therefore, the significant effects are those denoted as A, C, D, AC, and AD. The effect due to B is negligible.**

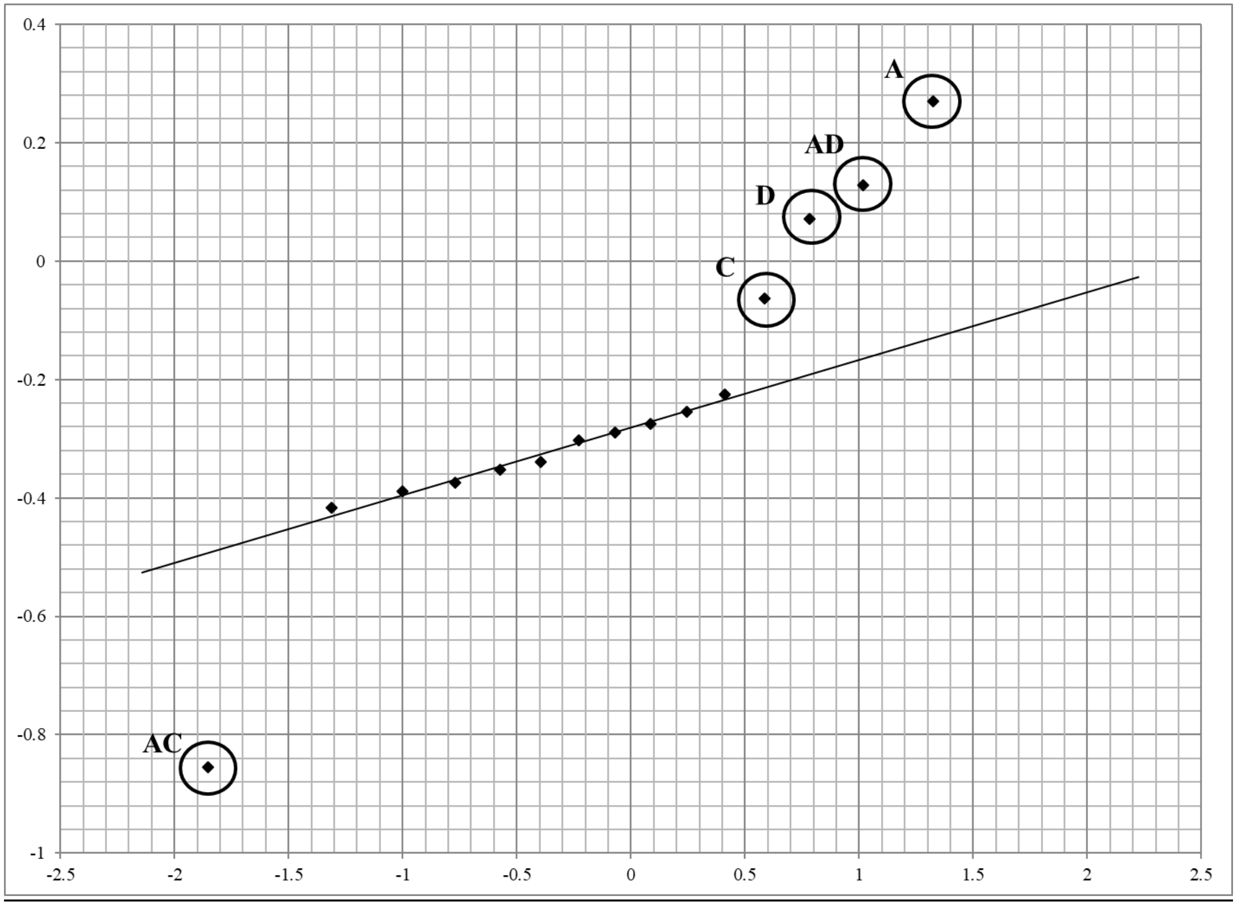


Figure 37: Normal probability plot of the effects

c) Dropping the B factor will produce a 2^3-factorial experiment with 2 replicates. In addition to dropping the terms associated with the B factor, all other terms will also be dropped. Since the design is orthogonal, we can drop the terms, without needing to recalculate anything. Therefore, the simplified model is given as

$$\underline{y = 70.06 + 10.8x_1 + 4.94x_3 + 7.31x_4 - 9.06x_1x_3 + 8.31x_1x_4}$$

d) The residuals for this case are shown in Figure 38. It can be seen that they are more or less normally distributed. Furthermore, since the reduced model has an **$R^2 = 0.966$** with all significant parameter values, it can be concluded that the results are probably good.

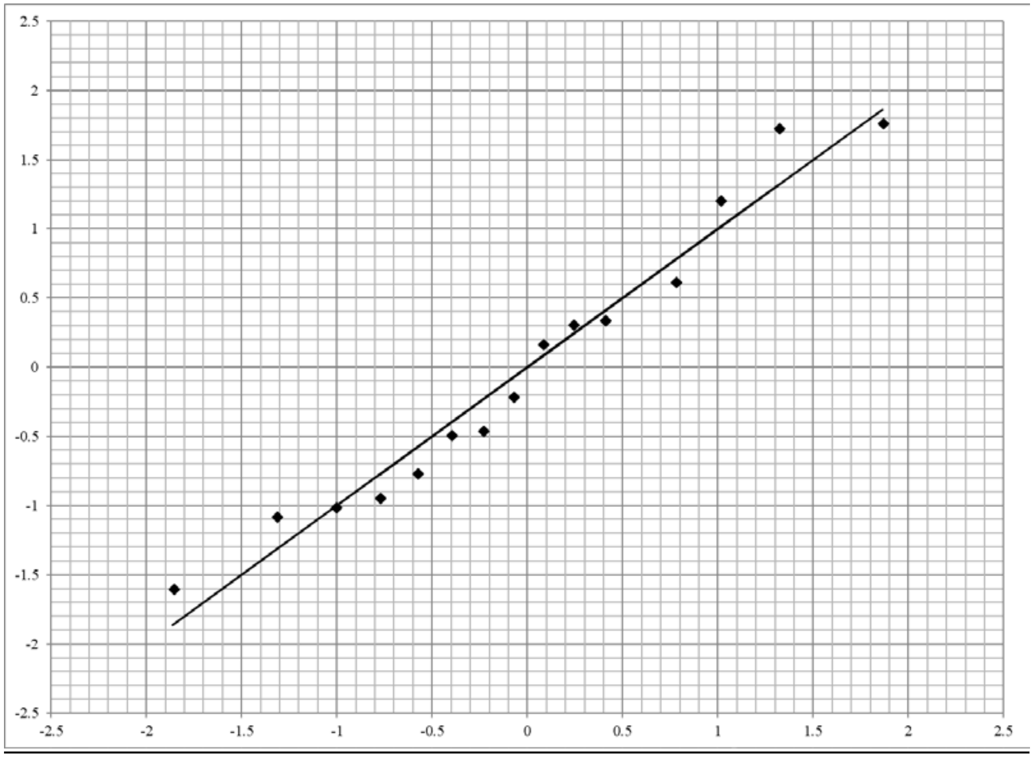
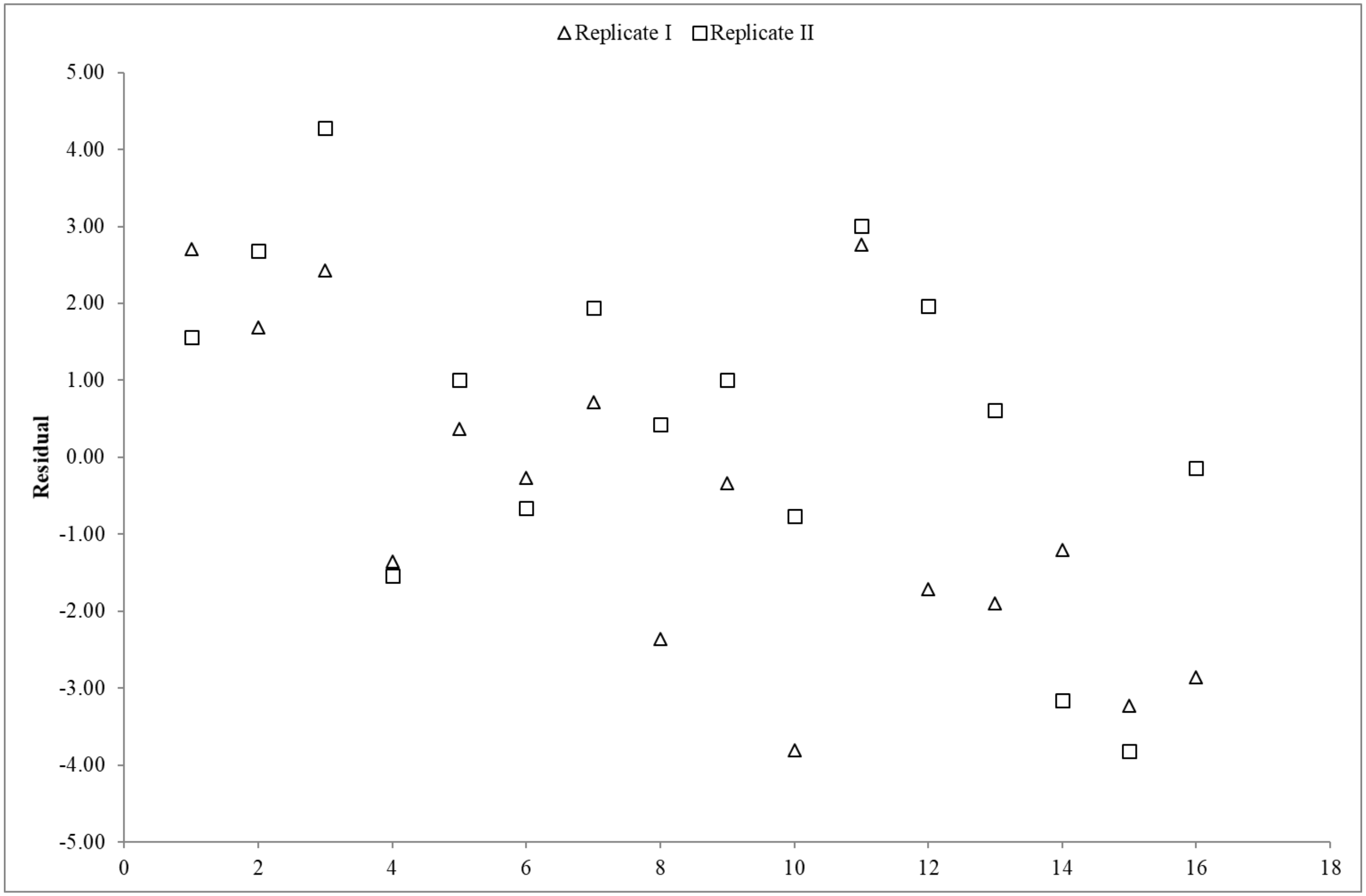
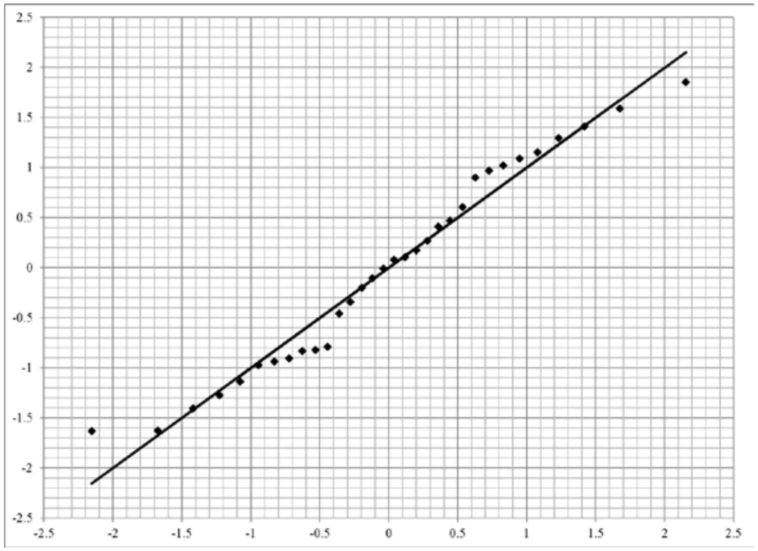


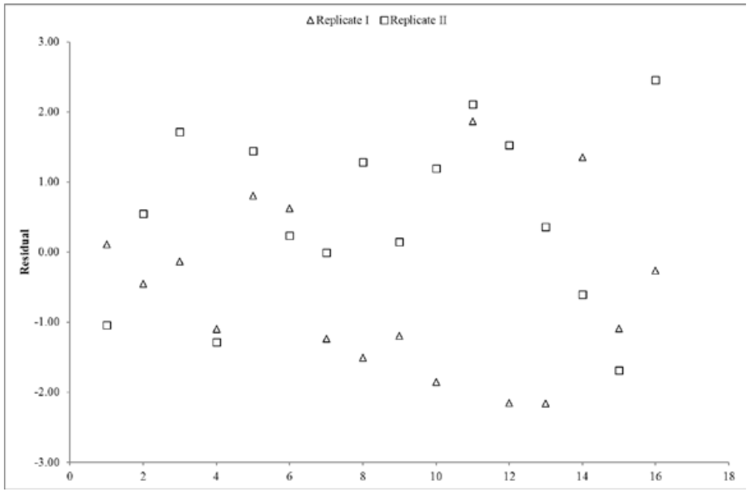
Figure 38: Normal probability plot of the residuals for the reduced model

Appendix III: Correction for Example 4.9, Figure 4.6 (bottom)



Appendix IV: Correction for Example 4.9, Figure 4.7





Appendix V: Updating Figures 4.10 and 4.11 for the Example in §4.7.4

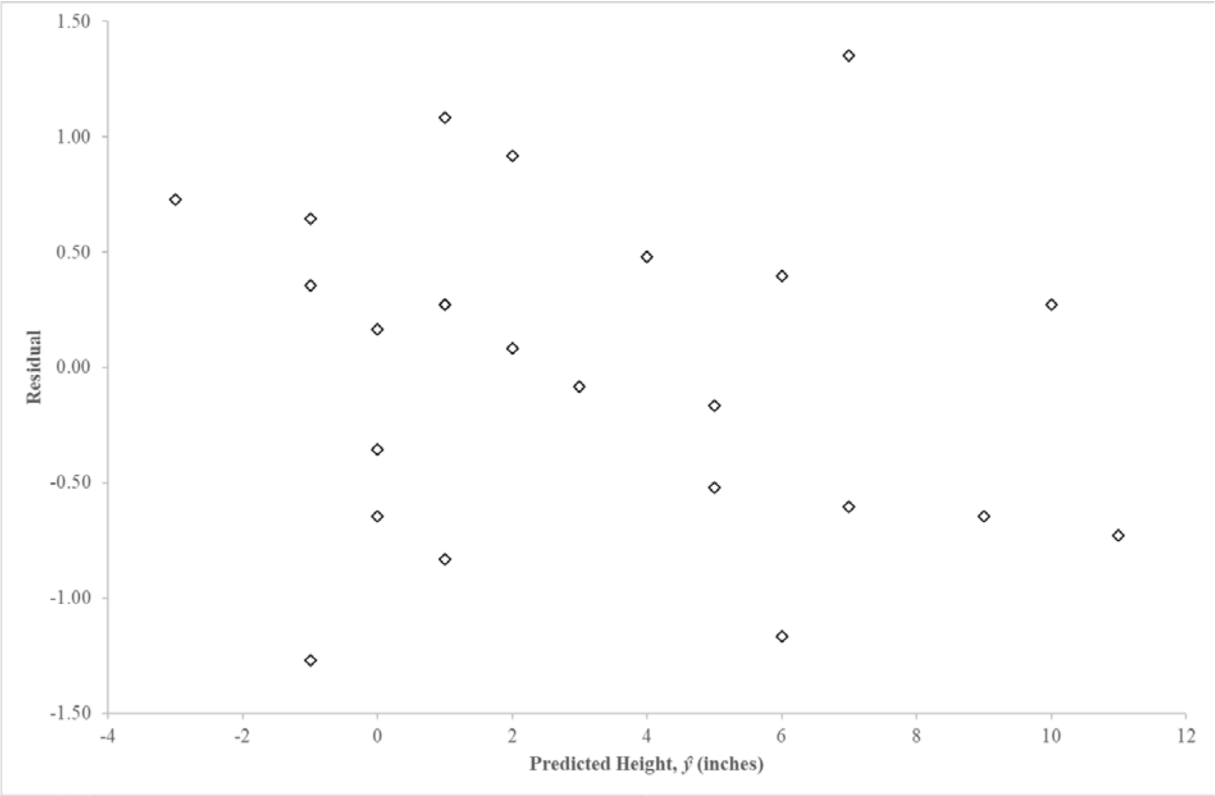


Figure 4.10: Residuals as a function of $\hat{y}$

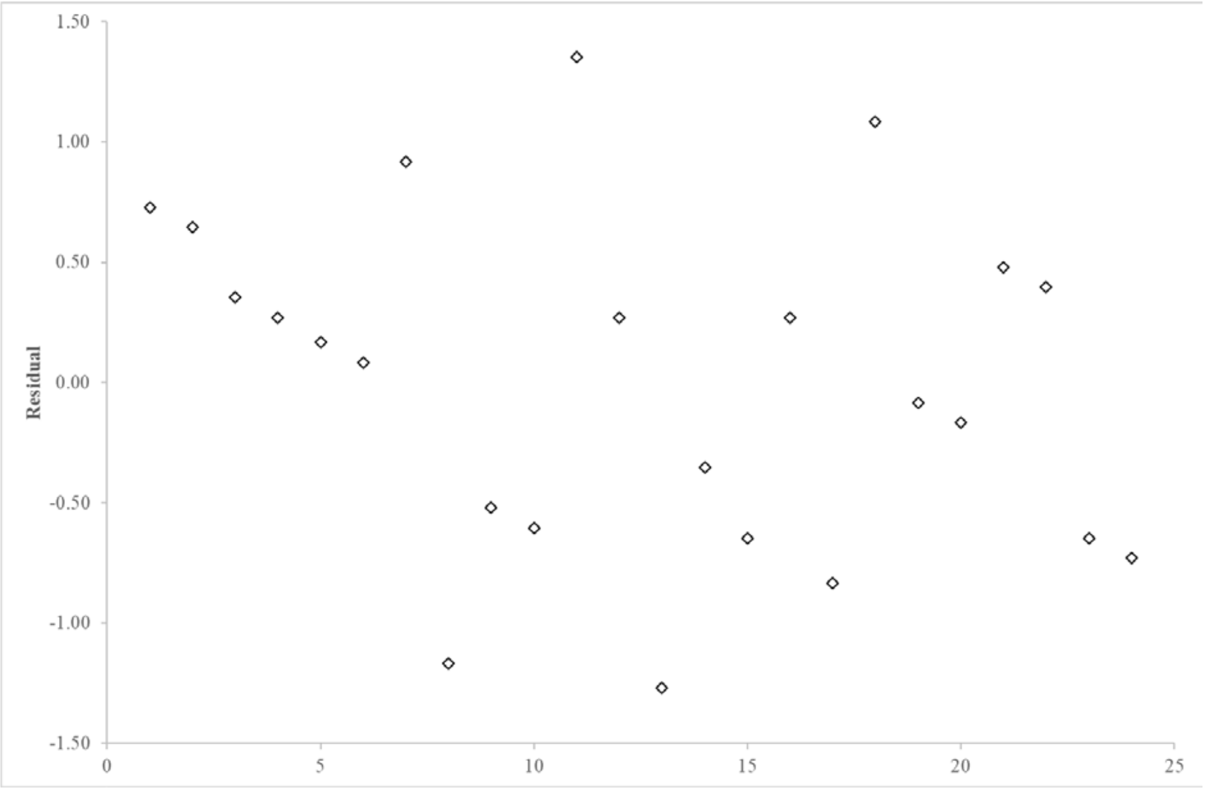


Figure 4.11: Time series plot of the residuals

Appendix VI: Updating Figures 4.13, 4.14, and 4.15 for the Example in §4.8.4

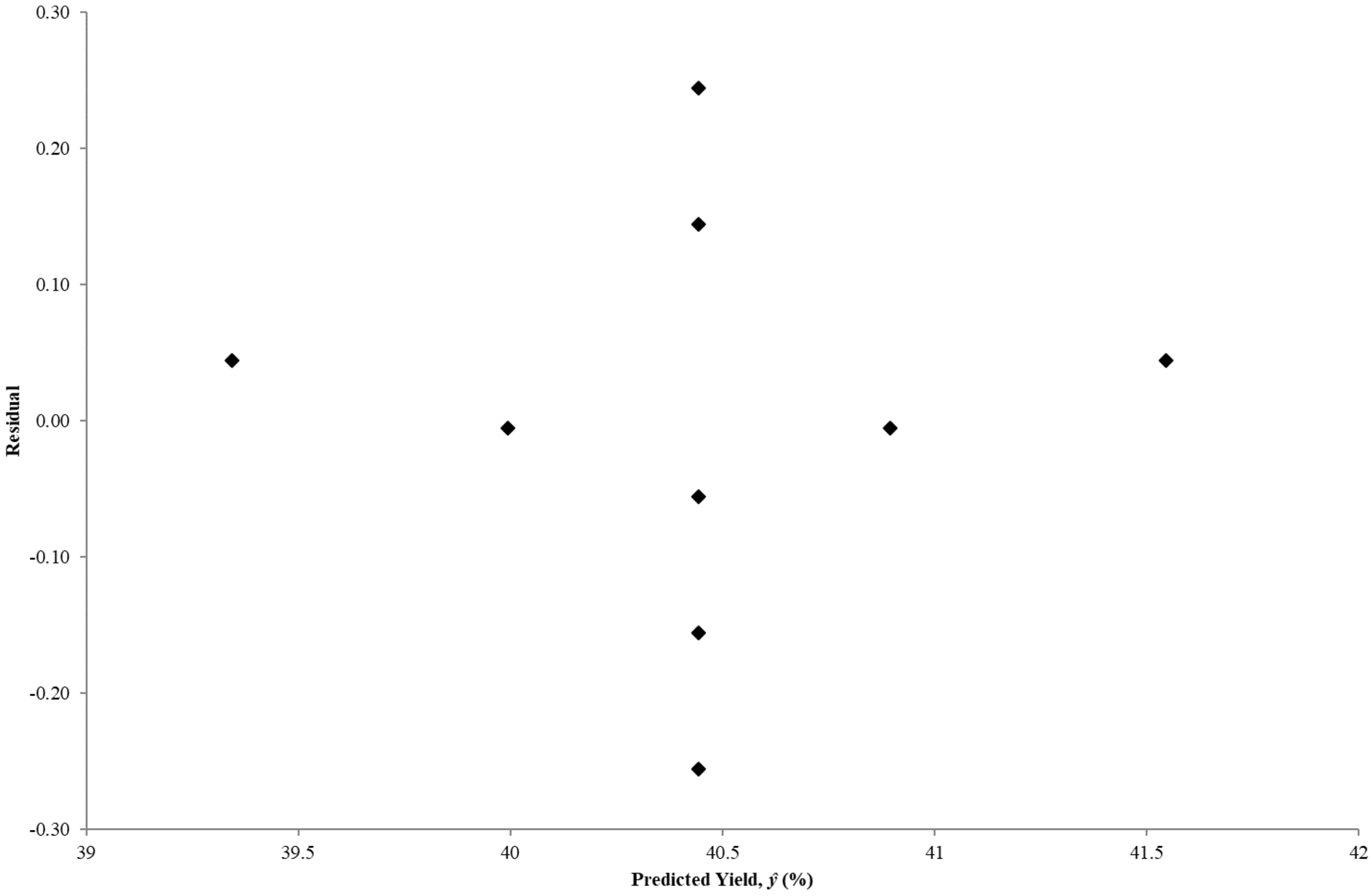


Figure 4.13

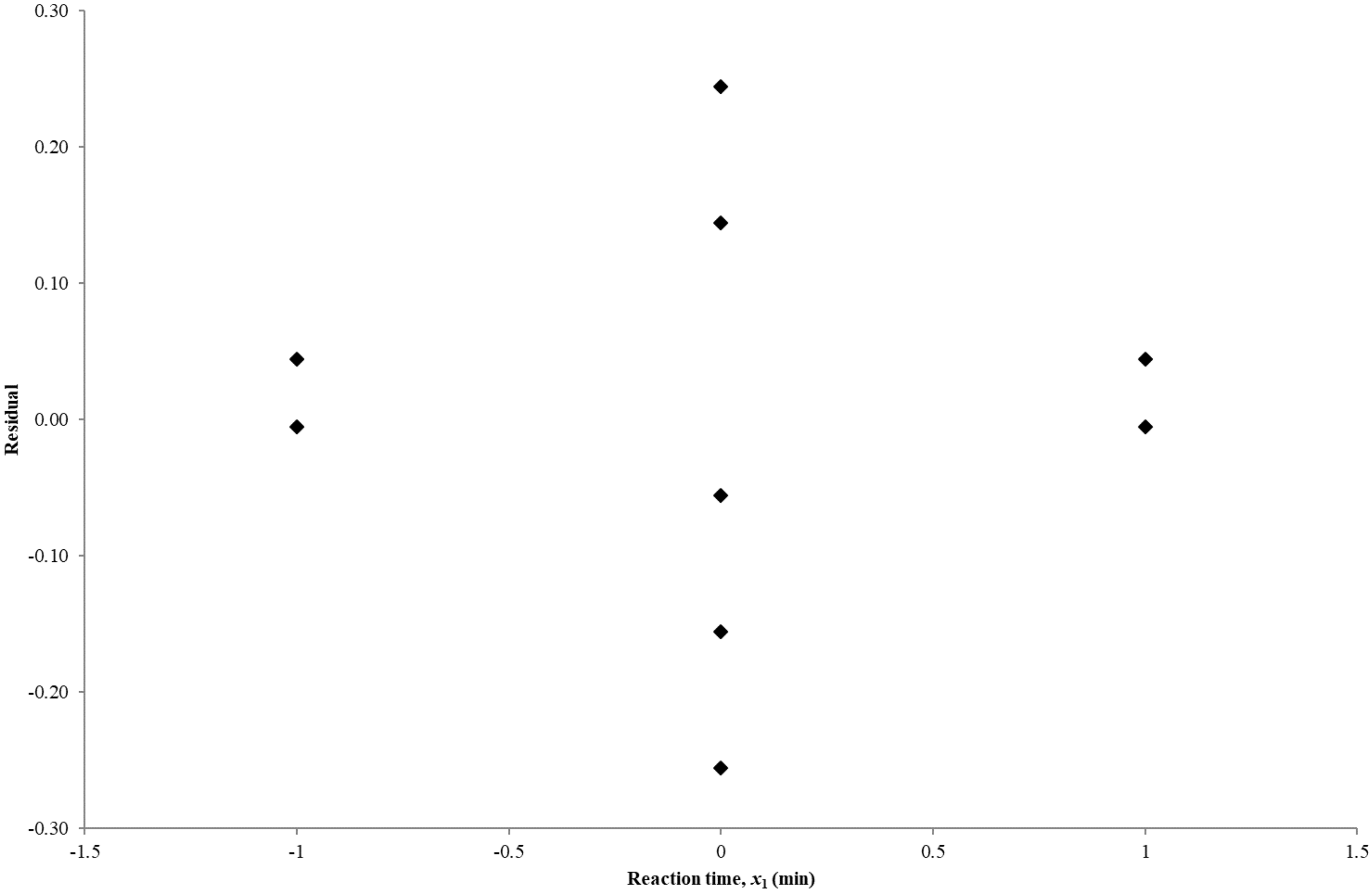

0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
Residual
-1.5
-1
-0.5
0
0.5
1
1.5
Reaction time, x_1 (min)

Figure 4.14

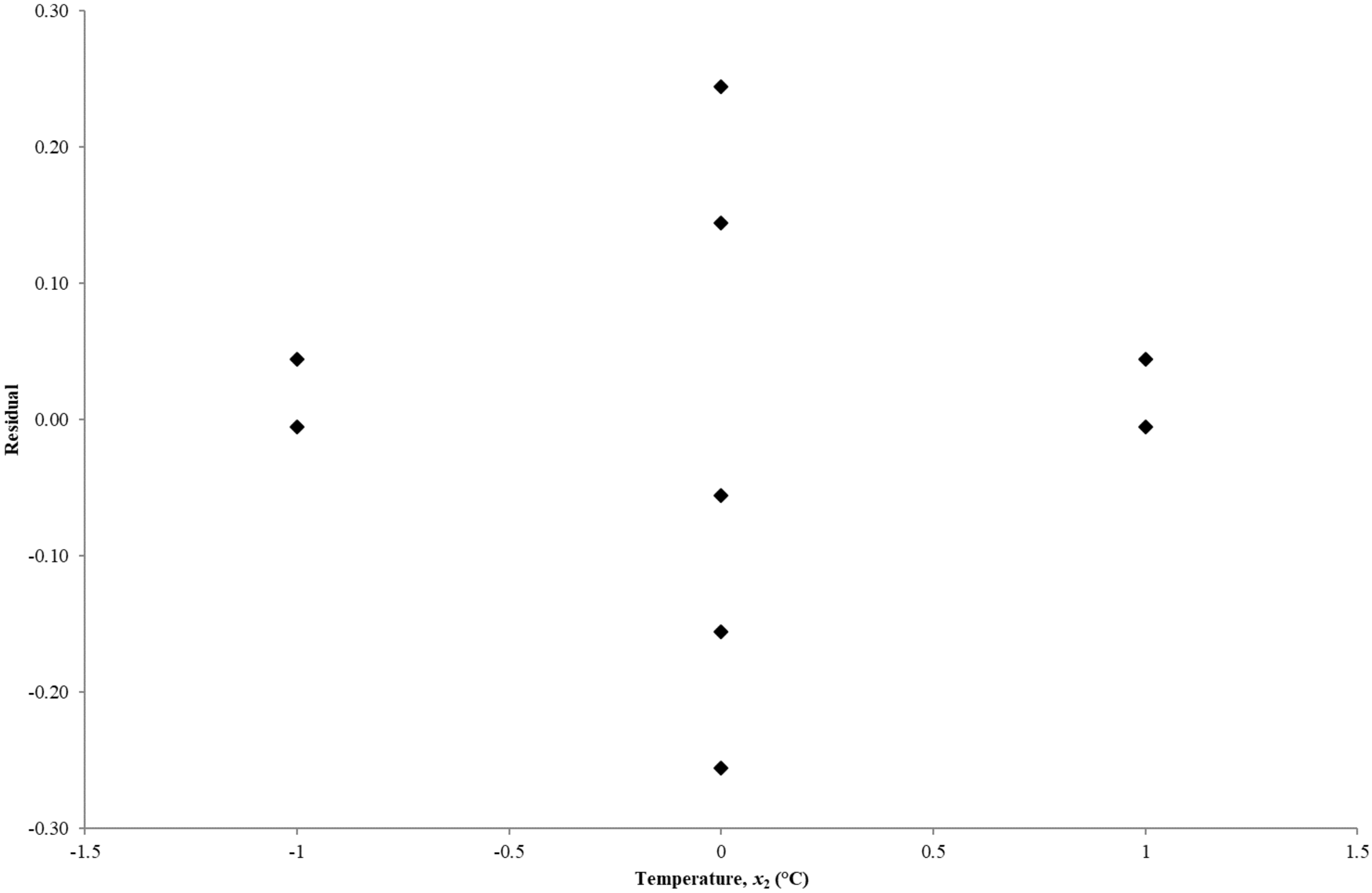

0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
Residual
-1.5
-1
-0.5
0
0.5
1
1.5
Temperature, x_2 (°C)

Figure 4.15

Appendix VII: Updating Text in §2.1

Let X be a **random variable** that assigns to each outcome in $\mathbb{S}$ a real number, that is, $X : \mathbb{S} \rightarrow \mathbb{R}$ for which $\{s : X(s) \leq x\} \in \mathbb{F}$, $\forall x \in \mathbb{R}$. Let the observed outcome at some given point be denoted by x. It should be noted that by convention random variables are denoted by capital letters, while the observations themselves are denoted by the corresponding lowercase letter. Thus, a random variable allows us to assign to each outcome in $\mathbb{S}$ a numeric value and, hence, compute various statistical properties, such as mean and variance. A random variable can be either discrete or continuous. A discrete random variable can only take certain values within a countable set (for example, flipping a coin), while a continuous random variable can take any values within (some subset of) $\mathbb{R}$. Often for continuous variables, the values assigned to the random variable are equal to the numeric values in $\mathbb{S}$. The process of obtaining an observation is called **sampling**. In the simplest case, the random variable can be viewed as assigning to a coin flip a payout based on the outcome, for example, we could assign \$1 for heads and −\$1 for tails. The individual observations would be determined based on the flips of the coin, for example, if heads was obtained, then $x_1 = \$1$.

Finally, the probability of obtaining a given observation can still be found using the probability measure function, P. Often, for discrete random variables, we write this probability as $P(X = x)$, where we mean the probability that the random variable X takes the specific value x. For continuous variables, we most commonly deal with intervals. The most common case is to seek the probability that the random variable will be less than or equal to a specific value x, that is, $P(X \leq x)$. For both discrete and continuous variables, other intervals can also be used, for example, $P(x_1 \leq X \leq x_2)$, where we wish to determine the probability that the value of the random variable will lie between two values x_1 and x_2.

Furthermore, it is now possible to define two functions $f_X(x)$ and $F_X(x)$ whose name and definition depend on whether we are dealing with discrete or continuous random variables. For a discrete random variable, we have that $f_X(x)$ is called the **probability mass function** and is defined as[7]

$$f_X\left(x\right) = \begin{cases} P\left(X = x\right) & \text{if } X\left(s\right) = x \text{ for some } s \in \mathbb{S} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

and $F_X(x)$ is called the **cumulative distribution function** (**cdf**) and is defined as

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} f_X\left(x\right) = \sum_{x_i \leq x} P\left(X = x_i\right) \quad (2)$$

For a continuous random variable, we first define the cumulative distribution function, $F_X(x)$, as

$$F_X(x) = P(X \leq x) \quad (3)$$

Then, the **probability density function**, $f_X(x)$, (**pdf**) is defined as

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (4)$$

In practice, it is the probability density function that is first defined and then integrated to obtain the cumulative distribution function, that is,

$$P(X \leq a) = \int_{-\infty}^{a} f(x)dx \quad (5)$$

It can be noted that the subscript on the probability density function is often not written unless it is desired to emphasis the associated random variable.

---

[7] Alternatively, we could define this function using indicator (or with a slight abuse of notation Dirac δ-) functions. This would then unify discrete and continuous variables when it comes to the computation of further properties.

By Kolmogorov’s axiom that $P(\mathbb{S}) = 1$, the probability density function has the following property,

$$\int_{-\infty}^{\infty} f(x)dx = 1. \qquad (6)$$

Furthermore, by Kolmogorov’s axiom that $P(\mathbb{E}) \geq 0$,

$$f(x) \geq 0 \text{ for all } x. \qquad (7)$$

The properties given by Equations (28) and (29) are useful for determining if a given candidate function is in fact a probability density function or if the result obtained is indeed correct.

In order to describe the resulting space for the random variable, it is useful to consider two terms previously introduced that will now be formally defined: **mean** and **variance**. For a discrete function, the mean, μ, is defined as

$$\mu = \sum_{\substack{\forall x: X(s)=x \\ \text{for some } s \in \mathbb{S}}} xP\left(X = x\right). \qquad (8)$$

The **variance**, σ^2, is defined as

$$\sigma^2 = \sum_{\substack{\forall x: X(s)=x \\ \text{for some } s \in \mathbb{S}}} \left(x - \mu\right)^2 P\left(X = x\right). \qquad (9)$$

The standard deviation would be defined as the square root of the variance.

For continuous functions, the mean would be computed as

$$\mu = \int_{-\infty}^{\infty} xf(x)dx \qquad (10)$$

while the variance would be computed as

$$\begin{aligned}\operatorname{var}(x) = \sigma^2 &= \int_{-\infty}^{\infty} (x-\mu)^2 \, f(x)dx \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2\end{aligned} \tag{11}$$

The variance can either be denoted by σ^2 or by *var*. It is common to use *var* when it is desired to treat the variance as an operator and perform additional manipulations with it.

Finally, the i^{th} **uncentred moment** of the probability density function $f(x)$, written as m_i, is

$$m_i = \int_{-\infty}^{\infty} x^i f(x)dx \tag{12}$$

It can be noted that the first moment is equivalent to the mean. In certain cases, centred moments are preferred. In such cases, the i^{th} **centred moment** for the probability density function $f(x)$, written as $\overline{m}_i$ is

$$\overline{m}_i = \int_{-\infty}^{\infty} (x-\mu)^i \, f(x)dx \tag{13}$$

The second centred moment, $\overline{m}_2$, is equivalent to the variance.

For a random variable, we denote the underlying probability space using a tilde, for example, $X \sim \mathfrak{N}(0,1)$ means that the random variable X is based on a normal distribution with a mean of zero and a variance of 1. Formally, we can state this as

$$P(X \le a) = \int_{-\infty}^{a} f_N(x)dx \tag{14}$$

where f_N is the probability density function for the normal distribution.