Statistics and Finance: An Introduction

David Ruppert

Errata: last updated March 9, 2009

Note: line $-n$ means the nth line from the bottom of the page.

Errors in 2nd printing

1. Page 7, line -4:

   "$A \in S$" should be "$A \subset S$".

2. Page 143, Result 5.4.1, second bullet:

   "(or smaller)" should be "(or larger)".

3. Page 182, line 1:

   "lower **ask price**" should be "higher **ask price**"

4. Page 331:

   "The third bootstrap sample is 82, 82, 93, 99, 104, 110. For this bootstrap sample $\overline{X}_{\text{boot}} = 95.00$, $\text{SE}_{\text{boot}} = 4.70$, and $t_{\text{boot}} = (98.5 - 95.00)/4.570 = 1.02$."

   should be

   "The third bootstrap sample is 82, 82, 93, 99, 104, 110. For this bootstrap sample $\overline{X}_{\text{boot}} = 95.00$, $\text{SE}_{\text{boot}} = 4.70$, and $t_{\text{boot}} = (98.5 - 95.00)/4.70 = 1.02$."

5. Page 351, lines -5 to -3:

   "The estimate of VaR(0.05) changed from 325.7 when $m = 100$ to 300.9 when $m = 50$. In a more extreme case, VaR(0.005) changed from 390.2 to 352.4."

   should be

   "The estimate of VaR(0.01) changed from 325.7 when $m = 100$ to 307.9 when $m = 50$. In a more extreme case, VaR(0.005) changed from 390.2 to 363.1."

6. Page 383, line immediately before Section 12.12:

   "see Section 12.4" should be "see Section 12.12"

7. Page 383, equation (12.10):

   coefficient of $g(\epsilon_{t-1})$ should be α_i, not α_1.

Errors in 1st printing but corrected in 2nd printing – 2nd printing became available June 2006

1. Page 3, line -9: "and it not easy to derive" should be "and is not easy to derive"

2. Page 8:

   "The events $A_1, \ldots, A_n$ are **independent** if

   $$P\{A_1 \cap \cdots \cap A_n\} = P\{A_1\} \cdots P\{A_n\}.$$

   "

   should be

   "The events $A_1, \ldots, A_n$ are **independent** if for any $1 \leq i_1 < \cdots < i_k \leq n$

   $$P\{A_{i_1} \cap \cdots \cap A_{i_k}\} = P\{A_{i_1}\} \cdots P\{A_{i_k}\}.$$

   "

3. Page 15, Result 2.4.1:

   $$F_Y(y) = F_X\{b^{-1}(y-a)\}$$

   should be

   $$\begin{aligned} F_Y(y) &= F_X\{b^{-1}(y-a)\},\ b > 0, \\ &= 1 - F_X\{b^{-1}(y-a)\},\ b < 0. \end{aligned}$$

   Also,

   $$F_Y^{-1}(p) = a + bF_X^{-1}(p)$$

   should be

   $$\begin{aligned} F_Y^{-1}(p) &= a + bF_X^{-1}(p),\ b > 0 \\ &= a + bF_X^{-1}(1-p),\ b < 0. \end{aligned}$$

4. Page 17, line -10:

   θX has density $f(\theta^{-1}x)$ should be θX has density $\theta^{-1}f(\theta^{-1}x)$.

5. Page 19, lines 4 and -7: "cdfn" is not available in the latest version of MATLAB, MATLAB 7. Instead use "cdf('norm',x,a,b)" to compute the CDF at "x" or the normal distribution with mean "a" and variance "b". Other CDF's can be computed as well. Type "help cdf" within MATLAB's command window.

6. Page 26, line 4:

   "Because skewness and kurtosis measure shape, they do not depend on the values of location and shape parameters." should be "Because skewness and kurtosis measure shape, they do not depend on the values of location and scale parameters."

7. Page 32, line 13:

   "the kurtosis of X is $190.2/3.84^2 = 12.9$" should be "the kurtosis of X is $190.2/3.4^2 = 16.45$"

8. Page 42, line −5:

$$E\{Y-(\beta_0+\beta_1X)\}^2=E(Y^2)-2\beta_0E(Y)-2\beta_1E(XY)+(\beta_0+\beta_1X)^2.$$

should be

$$E\{Y-(\beta_0+\beta_1X)\}^2=E(Y^2)-2\beta_0E(Y)-2\beta_1E(XY)+E(\beta_0+\beta_1X)^2.$$

9. Page 49, answer to Example 2.11: $E(0.3X_1+0.7X_2)=(.3)(1)+(0.7)(2)=1.35$ should be $E(0.3X_1+0.7X_2)=(.3)(1)+(0.7)(1.5)=1.35$.

10. Page 55, line 13: "The MAP estimate is $5/6=0.8333$" should be "The MAP estimate is $4/5=0.8$"

11. Page 58, last paragraph:

"Here is the output. Notice that the kurtosis is $\widehat{K}=4.08$, slightly larger than the value 3 of a normal distribution. Also, the p-values of the normality tests are small, indicating nonnormality. See Section 2.20 for a discussion of p-values and normality tests."

should be

"Here is the output. The excess kurtosis is $\widehat{K}=4.08$. The excess kurtosis is labeled simply "kurtosis" which can be misleading. The excess kurtosis is somewhat larger than the value 0 of a normal distribution. However, the p-values of the normality tests are large, so there is no reason to reject the null hypothesis of nonnormality. See Section 2.20 for a discussion of p-values and normality tests."

12. Page 59:

```
                    Tests for Normality

Test                   --Statistic---    -----p Value------

Shapiro-Wilk           W     0.994705    Pr < W      0.5311
Kolmogorov-Smirnov     D     0.041874    Pr > D     >0.1500
Cramer-von Mises       W-Sq  0.068111    Pr > W-Sq  >0.2500
Anderson-Darling       A-Sq  0.399967    Pr > A-Sq  >0.2500
```

should be

```
                    Tests for Normality

Test                   --Statistic---    -----p Value------

Kolmogorov-Smirnov     D     0.053255    Pr > D     <0.0100
Cramer-von Mises       W-Sq  2.336439    Pr > W-Sq  <0.0050
Anderson-Darling       A-Sq  13.96599    Pr > A-Sq  <0.0050
```

13. Page 64:

If the hypotheses are H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$, then H_0 is rejected if $|t| > t_{\alpha/2;n_1+n_2-2}$. If the hypotheses are H_0: $\mu_1 \leq \mu_2$ and H_1: $\mu_1 > \mu_2$, then H_0 is rejected if $t > t_{\alpha/2;n_1+n_2-2}$ and if they are H_0: $\mu_1 \geq \mu_2$ and H_1: $\mu_1 < \mu_2$, then H_0 is rejected if $t < -t_{\alpha/2;n_1+n_2-2}$.

should be

If the hypotheses are H_0: $\mu_1 - \mu_2 = \Delta_0$ and H_1: $\mu_1 - \mu_2 \neq \Delta_0$, then H_0 is rejected if $|t| > t_{\alpha/2;n_1+n_2-2}$. If the hypotheses are H_0: $\mu_1 - \mu_2 \leq \Delta_0$ and H_1: $\mu_1 - \mu_2 > \Delta_0$, then H_0 is rejected if $t > t_{\alpha;n_1+n_2-2}$ and if they are H_0: $\mu_1 - \mu_2 \geq \Delta_0$ and H_1: $\mu_1 - \mu_2 < \Delta_0$, then H_0 is rejected if $t < -t_{\alpha;n_1+n_2-2}$.

14. Page 73, problem 10: "5/6" should be "4/5".

15. Page 73, problem 11. (d) "$p_0 > 0$" should be $p_0 < 1$" and "$p < p_0$" should be "$p > p_0$".

16. Page 74, problem 16: "$\widehat{\sigma}_{1,2}$" should be "$s_{1,2}$".

17. Page 76, first two lines: "If $P_t = 2$ and $P_{t+1} = 2.1$, then $1 + R_t = 1.05$ or 105% and $R_t = 0.05$ or 5%." should be "If $P_t = 2$ and $P_{t+1} = 2.1$, then $1 + R_{t+1} = 1.05$ or 105% and $R_{t+1} = 0.05$ or 5%."

18. Page 81: See comment above about "cdfn" on page 19.

19. Page 82, line 6: "$\text{Var}(S_t|Z_0) = \sigma^2 t$" should be "$\text{Var}(S_t|S_0) = \sigma^2 t$"

20. Page 84, Figure 3.3: "GRM" should be "GRW" (six places)

21. Page 88, caption of Figure 2.5:

    Five independent geometric random walks and GE daily log returns. The geometric random walks have the same expected log return, volatility, and starting point as the GE log returns."

    should be

    "Five independent geometric random walks and GE daily prices. The geometric random walks have the same expected log return, volatility, and starting point as the GE prices."

22. Page 91, line 7: "call" should be "called"

23. Page 99, Problem 2: $r_1, r_2, \ldots$ should be $\ldots, r_{-1}, r_0, r_1, r_2, \ldots$

24. Page 99, Problem 5: Add "More specifically, suppose that $X_k = X_0 \exp(r_1 + \cdots + r_k)$ where X_0 is a fixed constant and $r_1, r_2, \ldots$ are iid $N(\mu, \sigma_2)$.

25. Page 106, line -7: "$Y_0 = \mu$" should be "$E(Y_0) = \mu$"

26. Page 107, Figure 4.4: "ρ" should be "ϕ" in the caption and four times in the figure. However, "$\rho(h)$" in the y-axis labels is correct.

27. Page 108, line 8: "$Y_t - \mu = \phi(Y_{t-1} - \mu)$" should be "$Y_t - \mu = \phi(Y_{t-1} - \mu) + \epsilon_t$"

28. Page 109, line −9:

$$\frac{\phi^{2t} - 1}{\phi - 1} \text{ should be } \frac{\phi^{2t} - 1}{\phi^2 - 1}$$

29. Page 114

"The SAS estimate of ϕ is -0.2254. SAS uses the model

$$Y_t = \beta_0 - \phi Y_{t-1} + \epsilon_t \tag{4.17}$$

so SAS's ϕ is the negative of ϕ as we define it."

should be

"The AUTOREG estimate of ϕ is -0.2254. AUTOREG uses the model

$$Y_t = \beta_0 - \phi Y_{t-1} + \epsilon_t \tag{4.17}$$

so AUTOREG's ϕ is the negative of ϕ as we define it."

30. Page 115, line 2: "$\beta_0 = (1 - \rho)\mu$" should be "$\beta_0 = (1 - \phi)\mu$".

31. Page 115, line 3: Add "Another difference is that ARIMA agrees with our notation about the signs of the coefficients."

32. Page 115, line −9: $p = 0.079$ should be $p = 0.0179$.

33. Page 121, paragraph below SAS code: Delete "(+0.533 in our notation)" and "(−0.078 in our notation)".

34. Page 126, line −3: "lag of 80" should be "lag of 60"

35. Page 128, equation (4.29):

$$\widehat{Y}_{n+2} = \widehat{\mu} + \widehat{\phi}(\widehat{Y}_{n+1} - \widehat{\mu}) = \widehat{\phi}\{\widehat{\phi}(Y_n - \widehat{\mu})\},$$

should be

$$\widehat{Y}_{n+2} = \widehat{\mu} + \widehat{\phi}(\widehat{Y}_{n+1} - \widehat{\mu}) = \widehat{\mu} + \widehat{\phi}\{\widehat{\phi}(Y_n - \widehat{\mu})\},$$

36. Page 129, line −13:

$$\widehat{Y}_{n+1} = \widehat{\mu} + \widehat{\phi}(Y_n - \widehat{\mu}) - \theta\,\widehat{\epsilon}_n$$

should be

$$\widehat{Y}_{n+1} = \widehat{\mu} + \widehat{\phi}(Y_n - \widehat{\mu}) - \widehat{\theta}\,\widehat{\epsilon}_n$$

37. Page 132, Fig. 4.8: The label on the y-axis should be "log return," not "log price."

38. Page 139, line 3: "0.4272" should be "0.4264".

39. Page 141, −3:

$$\widehat{\rho}_{12} = \frac{\widehat{\sigma}_{12}}{s_1 s_2}.$$

should be

$$\widehat{\rho}_{12} = \frac{\widehat{\sigma}_{12}}{s_{R_1} s_{R_2}}.$$

40. Page 142, last two lines: "The dotted line connecting F with R_1 mixes the risk-free asset with the first risky asset." should be "The dotted line connecting F with R_2 mixes the risk-free asset with the second risky asset."

41. Page 144, line 5:

$$\sigma_T = \sqrt{w_T^2\sigma_1^2 + (1-w_T)^2\sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2}\,.$$

should be

$$\sigma_T = \sqrt{w_T^2\sigma_1^2 + (1-w_T)^2\sigma_2^2 + 2w_T(1-w_T)\rho_{12}\sigma_1\sigma_2}\,.$$

42. Page 147, line −4 "a matrix" should be "a vector"

43. Page 149, eq (5.11) and line −6: "μ_p" should be "μ_P" (upper case "P")

44. Page 153:

"The variance of this portfolio is

$$\begin{aligned} \mathrm{Var}(\mu_P) &= (\boldsymbol{g} + \boldsymbol{h}\,\mu_P)^{\mathsf{T}}\boldsymbol{\Omega}(\boldsymbol{g} + \boldsymbol{h}\,\mu_P) \\ &= \boldsymbol{g}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{g} + 2\boldsymbol{g}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{h}\mu_P + \boldsymbol{h}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{h}\mu_P^2. \end{aligned} \tag{1}$$

To find the **minimum variance portfolio** we minimize this quantity over μ_P by solving

$$0 = \frac{d}{d\mu_p}\mathrm{Var}(\mu_P) = 2\boldsymbol{g}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{h} + 2\boldsymbol{h}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{h}\mu_P.$$

"

should be

"The variance of the return on the portfolio $R_P = \boldsymbol{\omega}_{\mu_P}^{\mathsf{T}}\boldsymbol{R}$ is

$$\begin{aligned} \mathrm{Var}(R_P) &= (\boldsymbol{g} + \boldsymbol{h}\,\mu_P)^{\mathsf{T}}\boldsymbol{\Omega}(\boldsymbol{g} + \boldsymbol{h}\,\mu_P) \\ &= \boldsymbol{g}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{g} + 2\boldsymbol{g}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{h}\mu_P + \boldsymbol{h}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{h}\mu_P^2. \end{aligned} \tag{2}$$

To find the **minimum variance portfolio** we minimize this quantity over μ_P by solving

$$0 = \frac{d}{d\mu_p}\mathrm{Var}(R_P) = 2\boldsymbol{g}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{h} + 2\boldsymbol{h}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{h}\mu_P.$$

”

Also

“Plugging $\mu_{\min}$ into (5.24) we find that the smallest possible variance of a portfolio is

$$\mathrm{Var}(\mu_{\min}) = \boldsymbol{g}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{g} - \frac{(\boldsymbol{g}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{h})^2}{\boldsymbol{h}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{h}}. \tag{3}$$

”

should be

“Plugging $\mu_{\min}$ into (5.24), and calling the portfolio $R_{\min}$, we find that the smallest possible variance of a portfolio is

$$\mathrm{Var}(R_{\min}) = \boldsymbol{g}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{g} - \frac{(\boldsymbol{g}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{h})^2}{\boldsymbol{h}^\mathsf{T}\boldsymbol{\Omega}\boldsymbol{h}}. \tag{4}$$

”

45. Page 158, line 9:

$$\mathbf{1}^\mathsf{T}\overline{\boldsymbol{\omega}} = \frac{V_1\,\sigma_2^2 + V_2\,\sigma_1^2 - (V_1+V_2)\,\rho_{12}\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2(1-\rho_{12})}.$$

should be

$$\mathbf{1}^\mathsf{T}\overline{\boldsymbol{\omega}} = \frac{V_1\,\sigma_2^2 + V_2\,\sigma_1^2 - (V_1+V_2)\,\rho_{12}\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)}.$$

46. Page 161, line −2: “(5.37)” should be “(5.36)”

47. Page 162:

```
bmu = [0.08;0.03;0.05] ;
bOmega = [ 0.3 0.02 0.01 ;
    0.02 0.15 0.03 ;
    0.01 0.03 0.18 ] ;
```

should be

```
bmu = [0.08;0.03;0.05] ;
bOmega = [ 0.3 0.02 0.01 ;
    0.02 0.15 0.03 ;
    0.01 0.03 0.18 ] ;
rf = .04 ;  % Risk-free rate
```

48. Page 167, line 3: “rate of return” should be “rate of expected return”

49. Page 167, line 8: "tangency return" should be "tangency portfolio".

50. Page 175, line 25: "The changes in the 30-year rate are the variable `cm10_dif`." should be "The changes in the 30-year rate are the variable `cm30_dif`."

51. Page 185, line 6: "SAS's PROC SAS" should be "SAS's PROC REG"

52. Page 198: The y-axis of the top, right plot should be labeled "RSTUDENT".

53. Page 210, footnote: "$h(y) = h(x)+h^{(1)}(x)(y-x)$" should be "$h(y) \approx h(x)+h^{(1)}(x)(y-x)$"

54. Page 226, line -16: "one million shares" should be "one billion shares".

55. Page 235, line 11: "then $\sigma_{\epsilon,P} = 0.0125\%$." should be "then $\sigma_{\epsilon,P} = 0.125\%$."

56. Page 236, line -2: "$\sigma_{j,M}$ is the correlation" should be "$\sigma_{j,M}$ is the covariance"

57. Page 252, last paragraph before Section 7.11: "If would be simple to add" should be "It would be simple to add".

58. Page 255, problem 6: "165." should be "165 $\%^2$. (Note: If returns are expressed in units of percent, then the units variance and covariances are percent-squared. A variance of $165\%^2$ equals 165/10,000.)". In part (c), "220" should be "220 $\%^2$".

59. Page 264, line 9: " \$20 at node F" should be " \$20 at node D"

60. Page 268:

$$E(Y_{t+1}|Y_t) = Y_t$$

should be

$$E(Y_{t+1}|Y_0, \ldots, Y_t) = Y_t$$

61. Page 270, line 8: "interest rate is 6%. then" should be "interest rate is 6%, then"

62. Page 273, add to end of Section 8.10.4: If $\sigma = 1$ then the Brownian motion is said to be standard.

63. Page 274, line 12: "$\exp(x) \approx 1 + x$" should be "$\exp(x) \approx 1 + x + x^2/2$"

64. Page 274, lines -9 and -6: "$2W_i - 1$" should be "$(2W_i - 1)$"

65. Page 274, line -2: "where B_t is Brownian motion" is "where B_t is standard Brownian motion"

66. Page 281, equation (8.23):

$$V_i = \beta_0 + \beta_1 K_i^c + \beta_2 (K_i^c)^2 + \beta_3 T_i^c + \beta_5 (T_i^c)^2 + \beta_6 (K_i^c)^3 + \epsilon_i.$$

should be

$$V_i = \beta_0 + \beta_1 K_i^c + \beta_2 (K_i^c)^2 + \beta_3 T_i^c + \beta_5 (T_i^c)^2 + \beta_6 (T_i^c)^3 + \epsilon_i.$$

(change in variable multiplying β_6)

67. Page 282, line 5:

$$\widehat{V} = \widehat{\beta}_0 + \widehat{\beta}_1 (K - \overline{K}) + \widehat{\beta}_2 (K - \overline{K})^2 + \widehat{\beta}_6 (K - \overline{K})^3.$$

should be

$$\widehat{V} = \widehat{\beta}_0 + \widehat{\beta}_1 (K - \overline{K}) + \widehat{\beta}_2 (K - \overline{K})^2.$$

68. Page 286, line −6:

"At node 1, the American option is worth

$$e^{-.05}\{(q)(4.91) + (1-q)(30)\} = 13.65,"$$

should be "At node 1, the American option is worth

$$e^{-.05}\{(q)(4.91) + (1-q)(30)\} = 13.54,"$$

69. Page 298, problem 1: "The exercise price is \$55." should be "The exercise price of a European call is \$55."

70. Page 299, problem 6. (a): "given S_1, S_1, S_2" should be "given S_0, S_1, S_2."

71. Page 299, problem 6. (d): Add the sentence "Your answer should be a function of r."

72. Page 300, problem 9:

$$\Gamma = \frac{\phi(d_1)}{S\sigma\sqrt{T}}.$$

should be

$$\Gamma = \frac{\phi(d_1)}{S\sigma\sqrt{T-t}}.$$

73. Page 317: In (9.29) and (9.30), "=" should be "≈".

74. Page 324, Problem 1. (b): "par \$1000 bond" should be "par \$1000 zero-coupon bond".

75. Page 324, Problem 3: "a par \$1000 coupon bond" should be "a par \$1000 20-year coupon bond",

76. Page 326:

$$\omega_i = \frac{T_i \,\mathrm{NPV}_i}{\sum_{j=1}^{N} T_j \,\mathrm{NPV}_j}$$

should be

$$\omega_i = \frac{\mathrm{NPV}_i}{\sum_{j=1}^{N} \mathrm{NPV}_j}$$

77. Page 346, last paragraph is Section 11.1: "VaR uses two parameters, the horizon and the confidence level, which are denoted by $1-\alpha$ and T, respectively." should be "VaR uses two parameters, the horizon and the confidence level, which are denoted by T and $1-\alpha$, respectively.

78. Page 348, 1st paragraph in Example 11.2: "$s = 0.0151$" should be "$s = 0.0141$" and "$-3.107 \times 10^{-4} - (1.645)(0.0151) = -0.0236$" should be "$-3.107 \times 10^{-4} - (1.645)(0.0141) = -0.0236$". Also, "VaR(0.05) is \$471 = (0.0326) × (\$20,000)." should be "VaR(0.05) is \$471 = (0.0236) × (\$20,000)."

79. Page 351, lines 8, 17: $\{(\log(k/n), -R_{(k)})\}_{k=1}^{m}$ should be $\left[\{\log(k/n), \log(-R_{(k)})\}\right]_{k=1}^{m}$

80. Page 352, Figure 11.3: Add to caption: " Dots: $-R_{(k)}$ is plotted on the x-axis, and k/n on the y-axis, both with log scales. Solid line: least-squares fit of $\log(k/n)$ (dependent of Y-variable) regressed on $\log(-R_{(k)})$ (independent or X-variable)."

81. Page 355, lines 3,4: "In (11.10), the holding period and confidence coefficient for the derivative are the same as for the asset." should be "In (11.10), the holding period, S, and confidence coefficient for the derivative are the same as for the asset."

82. Page 355, line −3: "which is denoted by" should be "which are denoted by"

83. Page 357, line 3: "a large confidence" should be "a large α (small confidence coefficient)"

84. Page 361, problem 4. (b) "Find the parametric estimate ..." should be "Find the nonparametric estimate ..."

85. Page 366, line −4: "taking expectations in (12.5)" should be "taking expectations in (12.6)"