Subject CS2

Corrections to 2019 study material

0 Comment

This document contains details of any errors and ambiguities in the Subject CS2 study materials for the 2019 exams that have been brought to our attention. We will incorporate these changes in the study material each year. We are always happy to receive feedback from students, particularly details concerning any errors, contradictions or unclear statements in the courses. If you have any such comments on this course please email them to CS2@bpp.com.

This document was last updated on **27 August 2019**.

1 Paper A Course Notes

Chapter 5

Correction added on 24 June 2019

Page 36

In the formulae for the backward integral equations, the summation should be over all values of $k \neq i$ (rather than $l \neq i$ as shown).

Chapter 6

Correction added on 16 November 2018

Page 26

There are some errors in the Core Reading at the bottom of page 26. It should say:

For example, if the force of mortality is governed by Gompertz law with shape parameter equal to 0.01 and rate parameter equal to 0.001, the force of mortality or hazard at age 30 can be calculated using:

```
hgompertz(30, shape = 0.01, rate = 0.001)
```

to be 0.00135.

Chapter 11

Correction added on 16 November 2018

Page 11

In the third paragraph, the reference for penalisation should be Chapter 12, not Chapter 17.

Page 14

Correction added on 8 February 2019

In the last line, it should say:

The value of w_x should be proportional to $\dfrac{E_x^c}{\hat{\mu}_x}$.

Page 21

Correction added on 16 November 2018

At the end of Section 5.1, it should say $\boldsymbol{D_x - E_x^c \mu_x^s}$ and not $\boldsymbol{D_x - E_x^c \mathring{\mu}_x}$.

Chapter 12

Corrections in blue added on 22 October 2018

The corrections in blue below apply only to Course Notes purchased before 16 November 2018.

Some formulae have mysteriously disappeared from pages 42 and 46 of Chapter 12. Replacement pages are given at the end of this document. (Replacements for pages 41 and 45 are also given so that you can print double-sided.)

Page 6

About halfway down the page, it should say:

'Similarly, for age 80 in calendar year **1997**, we have:'

Page 18

The second paragraph in the solution should read as follows:

'In the random walk model, ε_t is the error between the actual value of the *increment* (or increase) in the value of k_t and the predicted amount of that increment μ.'

Page 24

Correction added on 22 February 2019

There is a log missing from the LHS of the final equation on Page 24. It should say:

$$\frac{E(D_x)}{E_x^c}=\mu_x \quad\Rightarrow\quad \ln\left[\frac{E(D_x)}{E_x^c}\right]=\ln\mu_x \quad\Rightarrow\quad \ln E(D_x)=\ln E_x^c+\ln\mu_x$$

Page 26

In the paragraph before the italicised text, it should say:

'So, minimising $P(\theta)$ will attempt to select values of θ_j that minimise the sum **of the squares** of the 2nd differences …'.

Page 39

Correction added on 24 June 2019

The formula for the penalised log-likelihood should be:

$$l_p(\theta)=l(\theta)-\frac{1}{2}\lambda P(\theta)$$

ie the l's should be lower case (as they represent log-likelihoods).

Page 51

About halfway down the page after 'Similarly', there is a typo in the final expression in the equation for P_{HZD}. μ_{HD} should be μ_{HZ}, *ie* the equation should read:

$$P_{HZD} = \int_{t=0}^{1} p_{HH}(t)\,\mu_{HZ}\,p_{ZD}(1-t)\,dt = \mu_{HZ}\int_{t=0}^{1} p_{HH}(t)\,p_{ZD}(1-t)\,dt$$

Chapter 13

Corrections added on 18 March 2019

Page 31

In the solution to the first question, the first root should be $-\frac{1}{3}$. The minus sign is missing.

Page 69

In (ii)(b), the characteristic equation of the MA part has an incorrect sign. It should be:

$$1+\beta\lambda^2=0$$

Chapter 14

Correction added on 27 August 2019

Page 47

The example multivariate time series should be:

For example, the time series $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + e_t$ can be written as:

$$\begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ X_{t-2} \end{pmatrix} + \begin{pmatrix} e_t \\ 0 \end{pmatrix}$$

ie $\underline{X}_t = A\underline{X}_{t-1} + \underline{e}_t$

Chapter 16

Corrections added on 7 January 2019

Page 37

In Question 16.4, W should be defined as $X-u \mid X>u$ and not $X-u \mid X<u$.

Page 40

In the first paragraph of part (ii), it should say:

The distribution of $X-u \mid X>u$ is called the threshold exceedances distribution.

Chapter 17

Correction added on 18 March 2019

Page 13

In part (ii) of the solution, it should say:

$$
\begin{aligned}
C[u,v] &= F_{XY}(x,y) \\
&= \tfrac{1}{40}xy\left(x+4y\right) \\
&= \tfrac{1}{160}\left(-4+\sqrt{16+20u}\right)\left(-1+\sqrt{1+80v}\right)\left(-5+\sqrt{16+20u}+\sqrt{1+80v}\right)
\end{aligned}
$$

The formula in the second line of this equation comes from the solution on page 4.

Correction added on 8 February 2019

Page 18

In the block of equations in the middle of the page, the second line should say:

$$\min\left(P(X \le x), P(X \le y - 0.01)\right)$$

Corrections added on 27 August 2019

Pages 37 and 47

The symbols for the coefficients of upper and lower tail dependence in the table headings should be λ_U and λ_L.

Page 40

The first sentence in the paragraph above the table should read:

The following table is required for (i)(b).

Correction added on 18 March 2019

Page 41

The heading for part (i)(b) of the solution should be 'Gaussian copula with $\rho = 0$'.

Correction added on 4 April 2019

Page 52

The first bullet point in Solution 17.3(i) contains a typo. It should say:

$$\psi(0) = \lim_{t \to 0} \frac{1}{\alpha}\left(t^{-\alpha} - 1\right) = \frac{1}{\alpha} \lim_{t \to 0}\left(\frac{1}{t^{\alpha}} - 1\right) = \infty$$

Correction added on 18 March 2019

Page 56

The first two paragraphs of Solution Q17.7(ii) should be changed as follows:

In terms of increasing probabilities, the order is Frank (0.0583), Gumbel (0.0986), then Clayton (0.1089). Each of these is higher than 0.0198, which we would get if we used the product copula. This is because the Frank, Gumbel and Clayton copulas all assume that X and Y are positively correlated (whereas the product copula assumes they are independent).

Tail dependence also has an effect on probabilities of the form $P(X \le a, Y \le a)$, but this is only significant if we are considering very small values of a, *ie* deaths in the very near future.

Chapter 19

Correction added on 29 October 2018

Page 23

In the first line, the formula reference should be **(19.17)** and not (5.17).

Chapter 20

Correction added on 29 October 2018

Page 18

In (i), the references should be to formulae **(19.9) and (19.10)**, not (19.8) and (19.9).

Chapter 21

Corrections added on 16 November 2018

Page 29

The heading of Section 5.5 should be ‘Splitting the data into training, validation and testing data sets’.

Pages 43 and 44

To clarify:

- For a binary classification problem (*ie* one in which the data points are divided into two distinct categories), the Gini index must take a value between 0 (perfect purity, *ie* all the points at each node are of the same type) and 0.5 (worst purity).
- For a classification problem where the data points are divided into m distinct categories, the Gini index must take a value between 0 and $1-\dfrac{1}{m}$. As $m \to \infty$, the upper limit for the Gini index tends to 1.

2 PBOR

Poisson processes (Chapters 1 and 4)

Correction added on 22 November 2018

Solution to Question 1.1 part (x)

The standard deviation of the total annual claim amount for each claim type has been calculated incorrectly.

The correct code is:

```
sd.annual.motor=sqrt(365*Poiss.param.motor2*((m.p.motor[2])^2+
(m.p.motor[1])^2))
sd.annual.motor
```

```
[1] 11307.54
```

```
sd.annual.house=sqrt(365*Poiss.param.house2*((m.p.house[2])^2+
(m.p.house[1])^2))
sd.annual.house
```

```
[1] 37786.36
```

```
sd.annual.travel=sqrt(365*Poiss.param.travel2*((m.p.travel[2])^2+
(m.p.travel[1])^2))
sd.annual.travel
```

```
[1] 1116.073
```

Markov jump processes (Chapters 4 and 5)

Correction added on 22 November 2018

Solution to Question 5.1 part (iii)

The function at the start of part (iii)(a) should read:

```
calc.prob.matrix=function(s,t,h){
  temp.matrix=diag(3)
  for (j in 1:((t-s)/h)){
    temp.matrix=temp.matrix %*% (diag(3)+gen.matrix(s+h*(j-1))*h)
    Ph=temp.matrix
  }
  Ph
}
```

Hence, applying this function for part (iii)(a) gives:

```
calc.prob.matrix(25,25+10,1/12)
```

```
          [,1]      [,2]       [,3]
[1,] 0.5089156 0.4626218 0.02846263
[2,] 0.5085204 0.4629208 0.02855880
[3,] 0.5072905 0.4636936 0.02901591
```

The required probability is 0.5085204.

Similarly, applying the function for part (iii)(b) gives:

```
calc.prob.matrix(60,60+5,1/12)

          [,1]      [,2]       [,3]
[1,] 0.2717631 0.6502690 0.07796783
[2,] 0.2631221 0.6533405 0.08353736
[3,] 0.2373654 0.6335591 0.12907549
```

The required probability is 0.6502690.

Applying the function for part (iii)(c) gives:

```
calc.prob.matrix(50,50+7,1/12)

          [,1]      [,2]       [,3]
[1,] 0.3223396 0.6126956 0.06496485
[2,] 0.3202409 0.6134061 0.06635304
[3,] 0.3118890 0.6119035 0.07620748
```

The required probability is 0.3118890.

Estimating the lifetime distribution (Chapter 7)

Correction added on 27 August 2019

Survival models – Course Notes and Questions

Page 3

The last row of the table should not be included. Correspondingly, part (b) of parts (i) and (ii) should say for $j = 0,1,2,\ldots 9$.

Graduation Methods (Chapter 11)

Correction added on 27 August 2019

Graduation methods – Course Notes and Questions

Page 3

The penultimate line in the analyst's code given should not divide `Phixj` by 1000. The line should read:

```
term=term+b[j]*Phixj}
```

Time series (Chapters 13 and 14)

Corrections added on 3 January 2019

Time series, removing trends – Solution 13-14.5

Page 3

Underneath the statement of the null hypothesis, it should say:

The *p*-value for the first test is greater than 0.05, so there is **insufficient** evidence at the 5% level to reject the null hypothesis and we **conclude that *s* should be differenced at least once**.

Time series, forecasting – Solution 13-14.13(i)

Page 4

In the heading of part (i), it should say December **2014** (not December 2015).

Also, in the last paragraph, it should say:

Since we need to forecast the time series up to December **2014**, we will set `n.ahead=60` in the predict function.

Correction added on 27 August 2019

Time series the basics – Course Notes

Page 3

In the solution for generating sample observations of the time series W_t, the `sd` argument should not be part of the list given as the second argument. The solution should read:

```
Wt<-10+arima.sim(n=50,list(ar=0.8,ma=c(0.4,0.1)),sd=5^0.5)
```

Loss distributions (Chapter 15)

Corrections added on 11 February 2019

Fitting a distribution – Questions

Page 3

The second sentence of Q15.6 should say 'The file 'exp.txt' gives the claim size for **10,000** of these claims.'.

Fitting a distribution – Solutions

Solution 15.5

In Solution 15.5 part (i), the code for calculating `logL` should read:

```
logL <- dgamma(datavector,shape=alpha[i],rate=beta[j],log=TRUE)
```

This gives the maximum row as:

```
[24,]  2.0 0.0020 -1563.410
```

Hence we select $\alpha = \mathbf{2.0}$ and $\beta = 0.002$.

Therefore, in part (ii), the code for plotting the fitted gamma distribution should read:

```
lines(seq(0:5000),dgamma(seq(0:5000),shape=2.0,rate=0.002),
col="blue")
```

Summary sheet

Page 3

For consistency with the *Tables*, the top row of the 4th column of the table should read $X \sim \log N(\mu,\sigma^2)$ and not $X \sim \log N(\mu,\sigma)$.

Extreme value theory (Chapter 16)

Correction added on 22 February 2019

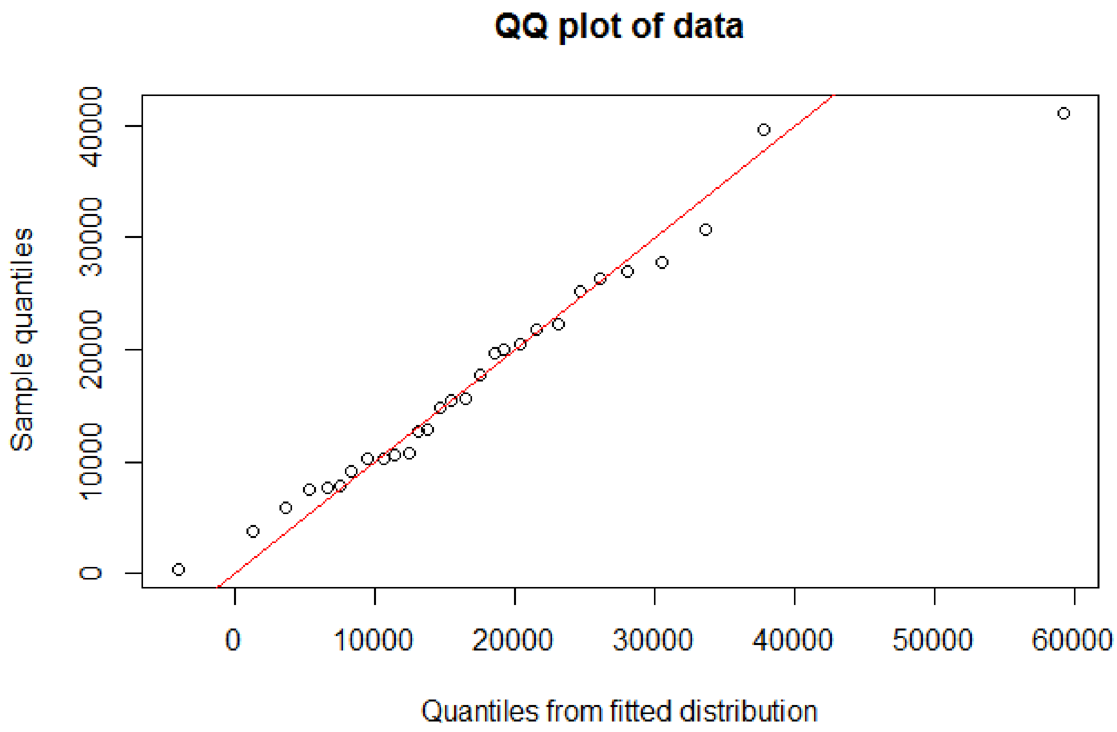
EVT and GPD distributions – Course Notes

Pages 10-11

On page 10, the calculation of `rp` should read:

```
rp<-b/g*((-log(runif(n)))^(-g)-1)+a
```

Hence the Q-Q plot on page 11 is incorrect. The corrected code gives a much better fit to the data:



Correction added on 22 February 2019

Chapter 16 – Summary sheet

Page 4

The code for simulating a GEV distribution is incorrect. The calculation of `rp` should read:

```
rp<-b/g*((-log(runif(n)))^(-g)-1)+a
```

Correction added on 27 August 2019

Measures of tail weight – Course Notes and Questions

Page 8

In Example 1, after the definition of the custom function to calculate `Sy`, the following paragraph and code should not be there and should be ignored:

The upper limit of the integration is infinity. However, we can't ask R to integrate over an infinite range, so instead we'll set the highest value of y to be high, in this case 10,000:

```
y<-c(1:10000)
```

Copulas (Chapter 17)

Correction added on 22 January 2019

Copulas – Course Notes and Questions

Page 7

The third paragraph should say:

We can now select all the rows where the **random number is less than the rescaled value of the copula density**. This will include (or exclude) them with the correct probability.

Machine learning (Chapter 21)

Correction added on 22 February 2019

Machine learning – Course Notes

Page 13

In the `fmax2` function, the calculation of `posterior` should be:

```
posterior=c(0.5,0.3,0.2)*c(f1(x),f2(x),f3(x));
```

3 Y assignments

Assignment Y1

Correction added on 25 March 2019

Question 2

There is a mistake in the solution to part (iv) of this question, which means that the 'Psh' matrix is updating 'too early' in the for loop. This solution should read as follows:

We can code a function to do this for us, using our functions created in parts (i) and (ii):

```
calc.prob.matrix=function(s,t,h){
      Psh=Ph(s,h)
      temp.matrix=diag(3)
      for (j in 1:((t-s)/h)){
            temp.matrix=temp.matrix%*%Psh
            Psh=Ph(s+h*j,h)
      }
      temp.matrix
}
```

4 ASET

September 2014 Solutions

Correction added on 25 March 2019

Solution 9

In the first box on page 42, it says $\sigma^2 = \text{var}(X_t) = \text{cov}(X_t, X_t) = \gamma_0$. This is potentially misleading as σ^2 denotes the variance of the white noise process (not the variance of X_t). For clarity, please delete '$\sigma^2 =$' from the start of this equation.

5 Flashcards

Chapter 13

Correction added on 25 March 2019.

In the solution to part (b) of Card 23, it should say:

To check invertibility, we examine the roots of the characteristic polynomial formed by the **moving average** part (see Card 12).

6 Revision booklets

Booklet 6

Correction added on 24 June 2019

Page 116

In the last line, it should say $Binomial(m,½)$ distribution, not $Binomial(\frac{1}{2}m,\frac{1}{4}m)$.

Chapter 12 Practice Questions

12.1 (i) Explain the notation and meaning of the parameters α_x and $f_{n,x}$ in the following reduction factor formula:

$$R_{x,t} = \alpha_x + (1-\alpha_x)(1-f_{n,x})^{t/n}$$

(ii) State briefly how the values of these parameters are usually determined.

(iii) The mortality rate for the base year of a mortality projection has been estimated to be:

$$m_{60,0} = 0.006$$

It is believed that the minimum possible mortality rate for lives aged 60 is 0.0012. It is also believed that 30% of the maximum possible reduction in mortality at this age will have occurred by ten years' time.

Using an appropriate reduction factor, calculate the projected mortality rate for lives aged 60 in 20 years' time.

(iv) Describe the advantages and disadvantages of using an expectation-based approach to mortality projections.

12.2 (i) Discuss a major difficulty that is present in a three-factor age-period-cohort mortality projection model that is not found in either an age-period or age-cohort model. [1]

Exam style

(ii) The following Lee-Carter model has been fitted to mortality data covering two age groups (centred on ages 60 and 70), and a 41-year time period from 1990 to 2030 inclusive:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$$

(a) Define in words the symbols a_x, b_x, k_t and $\varepsilon_{x,t}$.

(b) State the constraints that are normally imposed on b_x and k_t in order for the model to be uniquely specified.

(c) In this model k_t has been set to cover a 41-year time period from 1990 to 2030 inclusive, such that for projection (calendar) year t:

$$k_{t+1} = k_t - 0.02 + e_t$$

where e_t is a normally distributed random variable with zero mean and common variance.

Identify the numerical values of k_t ($t = 1990, 1991, \ldots 2029, 2030$), ignoring error terms. *Hint: they need to satisfy the constraint for k_t that you specified in part (b).* [5]

(iii) Mortality has been improving over time for both ages included in the model in part (ii). You have been given the following further information about the model:

$$b_{60} = 3b_{70}$$

$$\hat{m}_{60,2010} = 0.00176$$

$$\hat{m}_{70,2010} = 0.01328$$

where $\hat{m}_{x,t}$ is the predicted mortality rate at age x in calendar year t calculated from the fitted model (*ie* ignoring error terms).

(a) State what the above information indicates about the impact of the time trend on mortality at the two ages.

(b) Use the above information to complete the specification of the model.

(c) Use the model to calculate the projected values of $\hat{m}_{60,2025}$ and $\hat{m}_{70,2025}$. [6]

(iv) Describe the main disadvantages of the Lee-Carter model. [3]
[Total 15]

12.3 You have fitted a model to mortality data that are subdivided by age x and time period t, with a view to using the model to project future mortality rates. For a particular age x, the model is defined as:

$$\ln\left[E(D_{x,t})\right] = \ln E^c_{x,t} + a + bt + ct^2$$

where $D_{x,t}$ is the random number of deaths, and $E^c_{x,t}$ is the central exposed to risk for age group x in time period t ($t = 0$ is the year 1975).

(i) If $m_{x,t}$ is the central rate of mortality for exact age x in time period t, show that the above model is equivalent to:

$$m_{x,t} = A\,B^t C^{t^2}$$

stating the values of the parameters A, B and C.

Chapter 12 Solutions

12.1 (i) ***Interpretation of the reduction factor parameters***

α_x is the lowest level, expressed as a proportion of the current mortality rate at age x, to which the mortality rate at age x can reduce at any time in the future.

$f_{n,x}$ is the proportion of the maximum possible reduction (of $(1-\alpha_x)$) that is expected to have occurred by n years' time.

(ii) ***How the parameters are determined***

Both parameters could be set by expert opinion, perhaps assisted by some analysis of relevant recent observed mortality trends.

(iii) ***Projected mortality rate at age 60 in 20 years' time***

We can first calculate α_{60} as:

$$\alpha_{60} = \frac{0.0012}{0.006} = 0.2$$

We are also given that $f_{10,60} = 0.3$, so we need:

$$R_{60,20} = \alpha_{60} + (1-\alpha_{60})(1-f_{10,60})^{20/10} = 0.2 + 0.8 \times (1-0.3)^2 = 0.592$$

Hence the projected mortality rate at age 60 in 20 years' time is:

$$m_{60,20} = m_{60,0}\, R_{60,20} = 0.006 \times 0.592 = 0.003552$$

(iv) ***Advantages and disadvantages of using an expectation approach***

Advantages

- The method is easy to implement.

Disadvantages

- The effect of such factors as lifestyle changes and prevention of hitherto major causes of death are difficult to predict, as they have not occurred before, and experts may fail to judge the extent of the impact of these on future mortality adequately.
- Because the parameters are themselves target forecasts, there is a circularity in the theoretical basis of the projection model (because forecasts are being used to construct a model whose purpose should be to produce those forecasts).
- Setting the target levels leads to an under-estimation of the true level of uncertainty around the forecasts.

12.2 (i) ***Difficulty of age-period-cohort models***

Three-factor models have the logical problem that each factor is linearly dependent on the other two. So we need to ensure that the three arguments of the function work together in a consistent way in the formulae. [1]

(ii)(a) ***Definitions***

In the Lee-Carter model:

- a_x is the mean value of $\ln\left(m_{x,t}\right)$ averaged over all periods t [½]
- k_t is the effect of time t on mortality [½]
- b_x is the extent to which mortality is affected by the time trend at age x [½]
- $\varepsilon_{x,t}$ is the error term (independently and identically distributed with zero mean and common variance). [½]

(b) ***Constraints***

The constraints are:

- $\sum_t k_t = 0$ [½]
- $\sum_x b_x = 1$ [½]

(c) ***Numerical values of k_t***

k_t is a linear function of calendar year t, whose values must sum to zero over the 41-year time period. So the function needs to pass through zero when t takes its central value (2010). Hence:

$$k_t = -0.02 \times (t - 2{,}010)$$

which gives:

$$k_t = 0.4, 0.38, \ldots. -0.38, -0.4 \text{ for } t = 1990, 1991, \ldots 2029, 2030 \text{ respectively.}$$ [2]

(iii)(a) ***Effect of time trend at different ages***

Mortality rates at age 60 are assumed to be improving at three times the rate at which they are improving at age 70. [1]

(b) ***Complete the specification of the model***

We need values of a_x and b_x at both ages.

As mortality rates are improving at both ages, the values of b_{60} and b_{70} are both positive. Using $b_{60} + b_{70} = 1$ we have:

$$3b_{70} + b_{70} = 1 \;\Rightarrow\; b_{70} = 0.25, b_{60} = 0.75$$ [1]