Subject CS2

Corrections to 2020 study material

0 Comment

This document contains details of any errors and ambiguities in the Subject CS2 study materials for the 2020 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 **28 October 2020**.

1 Paper A Course Notes

Chapter 3

Correction added on 3 March 2020

Page 41

The formula in the solutions for part (ii) of Question 3.6 is incorrect. It should be:

$$\hat{q}_x = 1 - e^{-\hat{\mu}}$$

Chapter 12

Correction added on 16 January 2020

Page 18

The first sentence after the table containing $\hat{a}_x$ and $\hat{b}_x$ references the incorrect sum for the $\hat{b}_x$ values. It should say:

Note that the values of $\hat{b}_x$ in the table don't sum to 1 due to rounding.

Correction added on 6 January 2020

Page 23

The subscripts of the k terms in the derivation of $\text{var}(\tilde{\mu})$ are incorrect. The corrected derivation is as follows:

$$\text{var}(\tilde{\mu}) = \text{var}\left[\frac{\left(k_{t_n} - k_{t_n-1}\right) + \left(k_{t_n-1} - k_{t_n-2}\right) + \ldots + \left(k_2 - k_1\right)}{t_n - 1}\right]$$

$$= \sum_{i=2}^{t_n} \text{var}\left[\frac{\left(k_i - k_{i-1}\right)}{t_n - 1}\right]$$

$$= \sum_{i=2}^{t_n} \frac{\sigma^2}{\left(t_n - 1\right)^2}$$

$$= \frac{\sigma^2}{t_n - 1}$$

Correction added on 16 January 2020

Page 45

In question 12.2 part (c) (ii), the question should ask for the projected value of k_t in the year 2038 instead of 2030. The question should read:

Calculate the projected values of k_t for $t = 2023, 2028, 2033$ and 2038, ignoring error terms and given that $\hat{k}_{2018} = -0.4$.

Correction added on 7 August 2020

Page 51

In part (c), the value of 0.7 for the time trend component in the year 2033 is an estimate. So, the equation should be:

$$\hat{k}_{2033} = 0.7$$

Chapter 14

Corrections added on 5 June 2020

Page 16

The sentence after the formula given for $\hat{\rho}_k$ incorrectly references r_k twice and should be as follows:

The notation r_k is sometimes used instead of $\hat{\rho}_k$.

Page 28

The following sentence, just above the section on counting turning points, should not be there and should be ignored:

Here the Core Reading is using N rather than n to denote the number of recorded values of the time series.

Chapter 16

Corrections added on 5 June 2020

Page 9

The solution to the example question at the bottom of the page only has one part, so the '(i)' can be ignored.

Page 29

The final expression for the first hazard function shown under the graph is incorrect. It should be:

$$\frac{0.2}{1+\frac{5}{x}}$$

Chapter 17

Correction added on 5 June 2020

Page 36

The two formulae given for the Gaussian copula around halfway down the page are incorrect. It should be:

The formula defining the bivariate Gaussian copula is mathematically equivalent to the following integral form:

$$C[u,v]=\frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\Phi^{-1}(v)}\left[\int_{-\infty}^{\Phi^{-1}(u)}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left(s^2+t^2-2\rho st\right)\right\}ds\right]dt$$

This can be simplified further to:

$$C[u,v]=\int_0^v\Phi\left(\frac{\Phi^{-1}(u)-\rho\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right)dt$$

Chapter 20

Correction added on 7 August 2020

Page 23

The first sentence in the R box at the top of the page is incorrect. One value of lambda is generated for each simulation. The sentence should be:

For each of 10,000 simulations, the following R code simulates a value of lambda from the $Gamma(0.1,1)$ distribution.

Chapter 21

Correction added on 1 November 2019

Page 29

In the last paragraph, the following sentence should not be there and should be ignored:

This is the same type of inequality that we saw in Section Error! Reference source not found.

Correction added on 6 January 2020

Page 54

In the second sentence, it is the number of factors, not combinations that can be chosen. The sentence should read as follows:

The number of such factors can be chosen so that the model is sufficient to give reasonably accurate results, but not so large that it reflects a lot of the random noise contained in the data.

Corrections added on 5 June 2020

Page 58

The following summary section on generalisation error should not be there and can be ignored:

Generalisation error

An upper bound can be determined for the magnitude of out-of-sample errors. This shows that, with a large enough training set, the out-of-sample error can be made as small as desired.

Page 66

In part (i)(a) of Question 21.6, the second summation range incorrectly references i instead of k. It should be:

$$\sum_{k=1}^{m} p_k \sum_{j=1, j\neq k}^{m} p_j = 1 - \sum_{k=1}^{m} p_k^2$$

Page 75

In part (i)(a) of Solution 21.6, the second summation range incorrectly references i instead of k. It should be:

$$\sum_{k=1}^{m} p_k \sum_{j=1, j\neq k}^{m} p_j = \sum_{k=1}^{m} p_k \left(1 - p_k\right) = \sum_{k=1}^{m} \left(p_k - p_k^2\right) = \sum_{k=1}^{m} p_k - \sum_{k=1}^{m} p_k^2 = 1 - \sum_{k=1}^{m} p_k^2$$

2 PBOR

Poisson Processes (Chapters 1 and 4)

Correction added on 18 September 2020

Solutions

Page 3

At the bottom of page 3, the calculation for n needs to have a one subtracted (as the first row does not have a waiting time). It should be:

```
n = length(test.data) – 1
```

Page 4

In the middle of page 4, the commands for calculating the entries in the final row of the table should only subtract the sum of the columns for all rows except the last row. It should be:

```
test.data[(largest.time+1),2] = n-sum(test.data[1:largest.time,2])
test.data[(largest.time+1),1] = n-sum(test.data[1:largest.time,1])
```

Pages 4 and 5

The table at the bottom of page 4 and the top of page 5 is incorrect. It should be:

```
   [,1]       [,2]        [,3]
0    60  65.698658 0.494297801
1   133 113.673573 3.285818917
2    76  93.555787 3.294351495
3    84  76.998417 0.636664515
4    70  63.371348 0.693358045
5    62  52.155977 1.857980489
6    38  42.925487 0.565175194
7    29  35.328596 1.133674318
8    27  29.076191 0.148250846
9    17  23.930328 2.007053632
10   19  19.695173 0.024537267
11   17  16.209550 0.038545913
12   16  13.340807 0.530051095
13    5  10.979770 3.256684550
14    5   9.036586 1.803117397
15    7   7.437304 0.025712894
16    7   6.121060 0.126209313
17    7   5.037764 0.764301520
18    5   4.146188 0.175823100
19    5   3.412401 0.738620554
20    3   2.808479 0.013060502
21    2   2.311439 0.041962640
22    2   1.902364 0.005011065
23    3   1.565686 1.313965017
24    1   1.288593 0.064633277
25    2   1.060540 0.832204096
26    1   0.872847 0.018523176
27    5   4.059089 0.218106512
```

Page 5

The table in the middle of 5 is incorrect. It should be:

```
   [,1]       [,2]       [,3]
0    60  65.698658 0.49429780
1   133 113.673573 3.28581892
2    76  93.555787 3.29435149
3    84  76.998417 0.63666451
4    70  63.371348 0.69335805
5    62  52.155977 1.85798049
6    38  42.925487 0.56517519
7    29  35.328596 1.13367432
8    27  29.076191 0.14825085
9    17  23.930328 2.00705363
10   19  19.695173 0.02453727
11   17  16.209550 0.03854591
12   16  13.340807 0.53005110
13    5  10.979770 3.25668455
14    5   9.036586 1.80311740
15    7   7.437304 0.02571289
16    7   6.121060 0.12620931
17    7   5.037764 0.76430152
18   29  23.427625 1.32541643
```

The test statistic is also incorrect. It should be as follows:

```
        sum(test.claim.times[,3])
[1] 22.0112
```

The final paragraph should then be:

So, because the test statistic of **22.0112** is less than the critical value of 27.59, there is insufficient evidence to reject H0 and we can conclude that the waiting times are exponentially distributed.

Estimating the lifetime distribution (Chapter 7)

Correction added on 6 January 2020

Pages 6, 7 and 8

The Kaplan-Meier plot (page 7) and the Nelson-Aalen plot (page 8) have incorrect headers. The graphs show estimates, rather than estimators.

The code to produce corrected graphs is below (page 6 for the Kaplan-Meier code and page 8 for the Nelson-Aalen code). The corrected part is shown in bold:

```
plot(c(0,tj),c(1,SKM),type="s",
     xlim=c(0,25),ylim=c(0.5,1),
     main="Kaplan-Meier estimate of S(t)",
     xlab="Time t",ylab="Survival probability")

plot(c(0,tj),c(0,Lambda),xlim=c(0,25),ylim=c(0,0.35),type="s",
      main="Nelson-Aalen estimate of the integrated hazard",
      xlab="Time t",ylab="Integrated hazard")
```


Mortality projection (Chapter 12)

Correction added on 7 August 2020

Mortality projection – Course Notes

Page 4

The description of the impact of $f_{n,x}$ at the end of part (ii) is the wrong way around. It should be:

$f_{n,x}$ controls the rate at which the projected rates at age x will approach their long-term value. If $f_{n,x}$ is close to 1, the rates will converge very quickly, whereas if $f_{n,x}$ is close to 0 they will converge very slowly.

Time series (Chapters 13 and 14)

Correction added on 6 January 2020

Multivariate time series – Course Notes

Page 3

The superscripts of the error terms in the matrix form of the equations in Example 1 are incorrect. The equation should be:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0.2 & 0.4 \\ 0.1 & 0.3 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} e_t \\ e'_t \end{pmatrix}$$

Correction added on 18 March 2020

Multivariate time series – Course Notes

Page 5

The R code to create the matrix at the top of the page is correct; however, the output is incorrect. The output should be:

```
     [,1] [,2]
[1,] -0.6  0.5
[2,]  0.8 -0.1
```

Extreme value theory (Chapter 16)

Correction added on 8 September 2020

Measures of tail weight – Course Notes

Page 6

The second paragraph under the graph incorrectly mentions a relatively light tail instead of a relatively heavy tail. It should read:

Thereafter, the hazard rate decreases, corresponding to a relatively heavy tail, *ie* 'survival becomes easier' after *x* has reached about 400.

Risk models (Chapters 19 and 20)

Correction added on 6 February 2020

Parameter uncertainty – Questions

Page 3

Question 19-20.6 is not meant to be in this document and can be ignored. It is a duplicate of question 19-20.2 from the collective risk model questions document.

3 X Assignments

Assignment X2 Solutions

Correction added on 30 January 2020

Page 3

The final survival function expressions given for the Weibull and Exponential distributions in part (iii) are incorrect. They are each missing a t in the exponent and should be as follows:

Weibull: $S_X(t) = \exp\left(-\alpha t^{\beta}\right) = e^{-\alpha t}$ *when* $\beta = 1$

Exponential: $S_X(t) = e^{-\lambda t}$

Assignment X4 Solutions

Correction added on 28 January 2020

Page 16

The final sentence is incorrect and should be replaced with the following:

However, the mean residual lifetime of the $Gamma(2,1)$ distribution is always higher than that of the $Exp(1)$ distribution. So, the $Gamma(2,1)$ distribution has the heavier tail. [1]

Although the mean residual lifetime for the $Gamma(2,1)$ *distribution decreases, it does so asymptotically down towards 1. Therefore, the mean residual lifetime for the* $Gamma(2,1)$ *distribution is always higher than that of the* $Exp(1)$ *distribution.*

This result can also be shown by considering the limiting density ratio:

Let $X \sim Gamma(2,1)$ *and* $Y \sim Exp(1)$ *then*:

$$\lim_{x\to\infty} \frac{f_X(x)}{f_Y(x)} = \lim_{x\to\infty} \frac{xe^{-x}}{e^{-x}} = \lim_{x\to\infty} x = \infty$$

This indicates that the $Gamma(2,1)$ *distribution has the heavier tail.*

4 Y Assignments

Assignment Y1 Solutions

Correction added on 6 February 2020

Page 12

The total number of marks for the graph should be 5 instead of 4.

Page 13

The total number of marks for the graph should be 4 instead of 2.

Assignment Y2 Questions

Correction added on 3 March 2020

Page 2

The formula for the estimate of gamma just above part (ii) of Question Y2.1 should have Q3 instead of Q2 in the denominator:

$$\hat{\gamma} = \frac{\ln\left[\dfrac{\ln(0.75)}{\ln(0.25)}\right]}{\ln\left(\dfrac{Q1}{Q3}\right)}$$

Assignment Y2 Solutions

Corrections added on 3 March 2020

Page 1

The first line of R code in part (ii) of Solution Y2.1 should reference Q3 instead of Q2:

Using the formulae given in the question:

```
(g = log(log(0.75) / log(0.25)) / log(Q1 / Q3))
```

Pages 2 and 3

At the bottom of page 2 and the top of page 3, the code in part (iii) of Solution Y2.1 incorrectly references a dataset called `newdata`. The code should be as follows:

Let's quickly check that this function works, before going any further:

```
fMLE(Wparams, data$Data)
[1] 429.619
```


Now we use the `nlm()` function. This will return the 'scale' and 'shape' parameters that minimise the negative loglikelihood function:

```
nlm(fMLE,Wparams, data$Data)
```

Page 18

In part (iv)(a) of Solution Y2.4, the line of code that stores the `nlm` output near the top of page 18 references the incorrect data set. The code should be as follows:

```
MLE_values<-nlm(fMLE,GEV_params, x = blockmax$maximum)
```

5 Tutorial Handouts

Paper B preparation day handouts

Correction added on 18 March 2020

On page 43 of part 2 of the solutions (or page 120 of the combined solutions file), the function in Question 39 part (i) that calculates the negative log-likelihood is incorrect. It is missing a final line in the function to return the calculated value. It should be:

```
f.neglnL = function(params){
  ln.fx = dlnorm(x,params[1],sqrt(params[2]),log=TRUE)
  neglnL = sum(-ln.fx)
  neglnL
}
```

Correction added on 18 September 2020

On pages 41 and 42 of part 2 of the solutions (or pages 118 and 119 of the combined solutions file), the title of the graph in question 37 part (ii)(b) is incorrect. It should be:

```
plot(x, fx, type = "l", main = "X~Burr(0.1,200,5)", ylab = "f(x)")
```

6 ASET

April 2015 – Solutions

Correction added on 18 September 2020

Page 16

In the lightbulb box around halfway down the page, the first bullet point refers to graphical graduation. This is not part of the Subject CS2 syllabus and can be ignored.

April 2017 – Solutions

Correction added on 18 September 2020

Page 24

When carrying out the signs test, the solutions give the distribution for the number of positive deviations under the null hypothesis but then use the number of negative deviations to calculate the observed value of the test statistic.

Using the number of positive deviations, the observed value of the test statistic is:

$$z = \frac{\left(31 + \frac{1}{2}\right) - 40}{\sqrt{20}} = -1.901$$

The conclusion should then be:

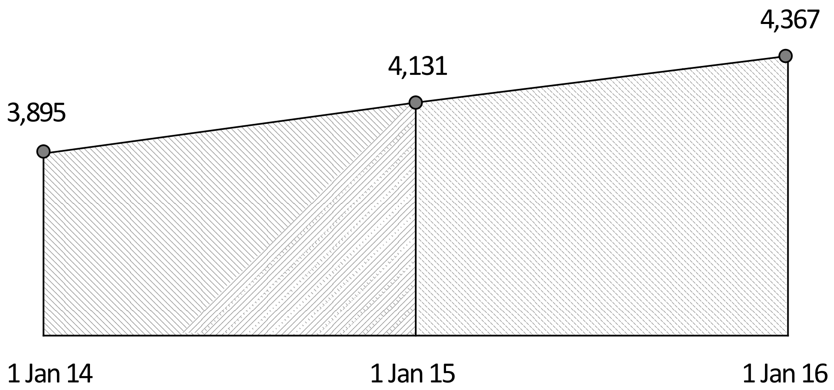
This is a two-sided test. Since –1.901 is between –1.96 and 1.96 (the lower and upper 2.5% points of the standard normal distribution), we have insufficient evidence to reject H_0 at the 5% significance level. We conclude that it is reasonable to assume that the graduated rates are equal to the underlying rates.

September 2016 – Solutions

Corrections added on 22 September 2020

Page 12

The chart at the top of page 12 is incorrect. The years are all off by one. The chart should be:



Page 13

The first paragraph uses the incorrect dates. The assumption should be:

We are also assuming that the number of policies in force varies linearly over the period from 1 January 2014 to 31 December 2015.

7 Mock exam

Solutions

Correction added on 28 October 2020

Page 17

At the end of part (v)(a), the general expression for ρ_k is for $k \geq 1$ not $\rho \geq 1$. It should read:

$$\rho_k = 0.5\rho_{k-1} = 0.5^{k-1}\rho_1 = 0.5^{k-1}\left(\frac{2}{19}\right) \quad k \geq 1$$