Exam Q1 (copy)

CEGM1000 Modelling, Uncertainty and Data for Engineers EXAM 24/25 · 5 exercises · 100.0 points

Exercise 1: Programming #43992451

10 pts · Last updated 24 Jan, 16:45

Programming part I: Object-oriented programming

Consider the following piece of code and output, and reflect on the object oriented programming (OOP) that is used:

Text

Screenshot 2024-10-28 153246.png

Image 1/10 Page

```
[ ]: import numpy as np
     data = np.genfromtxt(…)
     print(data.shape)
     print(data.mean())
```

```
Output:
(2734, )
65.4
```

Which of the following are OOP **attributes**? (You can select more than 1)

2 pts · Multiple choice · 8 alternatives

- [ ] data
- [ ] np
- [ ] numpy
- [ ] genfromtext
- [ ] print
- [x] shape
- [ ] mean
- [ ] none of the above

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

which of the following are OOP **methods**? (You can select more than 1)

2 pts · Multiple choice · 8 alternatives

- [ ] data
- [ ] np
- [ ] numpy
- [x] genfromtext
- [ ] print
- [ ] shape
- [x] mean
- [ ] none of the above

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

Programming part II: Errors

A friend has sent you a piece of code (not provided) that is supposed to compute the numerical solution to an ordinary differential equation using Taylor series, an iteration scheme and a numpy array. However, instead of getting the expected answer of 5, the code returns 9.

Text

What type of error is this? Add an explanation for your reasoning.

3 pts · Open · 3/10 Page

Logic error named correctly	1 pt
Reasoning given: the code executes without any syntax errors or exceptions but produces an incorrect result. This indicates that the logic used to compute the numerical solution is flawed, leading to an incorrect output (9 instead of the expected 5).	2 pts
Partially correct reasoning	1 pt

What would be a useful approach to help figure out and solve the problem? Consider each of the following and select **one** that you would find useful. Add an explanation for your selection.

- Read the traceback
- Use list comprehension
- Try pandas instead of numpy
- Use assert statements

3 pts · Open · 2/5 Page

Selecting assert statements	1 pt
Reasoning: Because they allow you to verify assumptions at specific points in the code. As there is no error message or traceback, the logic of the code needs to be checked. By placing assertions at critical steps (e.g., checking intermediate results or ensuring values remain within expected ranges), you can quickly identify where the computation diverges from expectations.	2 pts
Partially correct reasoning	1 pt

Exercise 2: Uncertainty Propagation #43992452

18 pts · Last updated 24 Jan, 16:45

Let X_1, X_2, X_3, be independent random variables, each following a standard normal distribution. These random variables are combined into a 3-dimensional random vector $Y = [Y_1, Y_2, Y_3]^T$, defined as:

Text

$$Y = \begin{bmatrix} X_1 \\ X_2 + X_3 \\ X_1 + X_2 + X_3 \end{bmatrix}$$

Text

What is the expectation of Y?

4 pts · Open · 4/5 Page

Linear propagation law, so just fill in expectation 0 for all the X_i

$E(Y) = [0, 0, 0]^T$ — 4 pts

Correct applying propagation law, but not knowing $E(X_i) = 0$ — 3 pts

(somewhat) correct approach, but some error(s) in the calculation, — 2 pts

Apply the propagation law to show that $cov\left(Y_{1,}Y_{2}\right)=0$ and **explain** why this is the case?

8 pts · Open · 9/10 Page

Model answer

$$Y=\begin{bmatrix}1&0&0\\0&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix}$$

$$\Sigma_Y=\begin{bmatrix}1&0&0\\0&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&1\\1&1&1\end{bmatrix}^T=\begin{bmatrix}1&0&1\\0&2&2\\1&2&3\end{bmatrix}$$

The covariance between the first and second element is zero because the second element contains X_2, X_3 which are independent of $Y_1 = X_1$.

$$Y=\begin{bmatrix}1&0&0\\0&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix}$$

According to the error propagation law:

8 pts

$$\Sigma_Y=\begin{bmatrix}1&0&0\\0&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&1\\0&1&1\end{bmatrix}^T=\begin{bmatrix}1&0&1\\0&2&2\\1&2&3\end{bmatrix}$$

The covariance between the first and second element is zero because the second element contains X_2, X_3 which are independent of $Y_1 = X_1$.

Correct implementation of propagation law, but not knowing standard normal distributions have a standard deviation of 1.	7 pts
correct explanation	3 pts
correct application of propagation law, explanation missing	5 pts
correct application propagation law, not using correct covariance matrix of X, and explanation missing	4 pts
some correct elements	2 pts

If the second element of Y is changed to $X_2 - X_3$, which elements in covariance matrix of the changed vector Y, Σ_Y, will be changed? Write out the new Σ_Y.

6 pts · Open · 9/10 Page

The covariance between the second and third element will be changed. The variance of the second element will not be changed.
The new matrix is shown below (not necessary to be computed), where the elements in red are changed: 6 pts

$$\Sigma_Y = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

somewhat correct 4 pts

a correct statement, but otherwise incorrect 2 pts

In between point descriptions 1 pt

Exercise 3: Observation theory #43992453

18 pts · Last updated 30 Oct, 23:24

An object is moving along a straight line with an unknown speed v. It starts at $t_0 = 0\,[s]$ with the known location $x_0 = 0\,[m]$. We measured its location every Δt seconds and obtained m measurements y_i, where $i = [1, \ldots, m]$, measured at $t_i = i \cdot \Delta t$.

Text

Model 1: $y_i = x_0 + vt_i$

Text

We assume the measurements y_i are independent and follow a normal distribution with standard deviation $\sigma_i = i \cdot \sigma$

Text

Specify the functional model and the covariance matrix of the observable, Σ_Y, according to Model 1.

6 pts · Open · 7/10 Page

$$E\begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix} = \begin{bmatrix} \Delta t \\ \vdots \\ m\Delta t \end{bmatrix} v$$

$$\Sigma_Y = \begin{bmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & m^2\sigma^2 \end{bmatrix}$$

6 pts

correct functional model	3 pts
correct stochastic model	3 pts
for error in notation or small error	-1 pts
for missing a square in covariance matrix	-1 pts
Larger error in functional or stochastic model, but there is still some merit to the solution	1 pt

Text

How many measurements m are required if we want the 99% confidence interval to be no larger than $\hat{v}\pm 0.2[m/s]$? Assume $\sigma = 0.4[m]$ and $\Delta t = 2[s]$

6 pts · Open · 9/10 Page

$$\hat{\sigma}_v = \frac{\sigma}{\Delta t\,\sqrt{m}}$$

$$2.576\,\frac{0.4}{2\,\sqrt{m}} < 0.2 \rightarrow m \geq 7$$

6 pts

Correct general formula for confidence interval

$$\pm k\sigma_{\hat{v}}$$

1 pt

correct k = 2.576 — 1.5 pts

correct approach for calculating

$$\sigma_{\hat{v}} = \sqrt{(A^T\Sigma_Y^{-1}A)^{-1}}$$

2 pts

correct answer — 1.5 pts

Another engineer assumes the initial position x_0 to be unknown, and that the movement may have an unknown acceleration a:

Text

Model 2: $y_i = x_0 + vt_i + \frac{1}{2}\,at_i^2$

Text

Which test should the engineer apply **and** what is the corresponding critical value used for significance level $\alpha = 0.05$? Answer the questions **without** specifying the functional model of Model 2.

6 pts · Open · 4/5 Page

Generalized Likelihood Ratio Test.
Since model 2 has two more unknowns than model 1, the test statistic follows a central Chi-square distribution with $q = 2$ degree of freedom, $\chi^2(2, 0)$

The critical value is 5.99155.9915

6 pts

correct mentioning of GLRT — 2 pts

correct q = 2 — 2 pts

correct critical value (for the q that the student used) — 2 pts

Exercise 4: Numerical Modeling #43992454

29 pts · Last updated 30 Oct, 23:42

Use the Newton-Raphson method to solve for x in the following equation:

$$x^5 + 3x = 100$$

Recall that:

$$x_{j+1} = x_j - \frac{g(x_j)}{g'(x_j)}$$

Where $g\left(x\right) = 0$ and j is the iteration number.

Write down the equation to be iterated in the Newton-Raphson method.
Assume a first guess of $x = 2$, then compute the second guess (i.e., one iteration).

8 pts · Open · 7/10 Page

This is the Newton-Raphson expression:

$$x^{j+1} = x^j - \frac{g(x^j)}{g'(x^j)}\,, \quad \text{where } g(x) = 0$$

So,

$$g(x) = x^5 + 3x - 100 \quad \text{and} \quad g'(x) = 5x^4 + 3.$$

The equation to be iterated is: 8 pts

$$x^1 = x^0 - \frac{g(x^0)}{g'(x^0)} = x - \frac{x^5 + 3x - 100}{5x^4 + 3}\,.$$

Substituting $x = 2$:

$$x^1 = 2 - \frac{2^5 + 3 \cdot 2 - 100}{5 \cdot 2^4 + 3} = 2 - \frac{32 + 6 - 100}{80 + 3} = 2.747.$$

Correct expression of g(x)	2 pts
Correct expression of g'(x)	2 pts
Correct implementation of the Newton-Raphson method	2 pts
Correct result	2 pts
correction	1 pt

What is the error order of the Taylor Series Expansion (TSE) of: $x^2 - a$ using the first 3 terms $\left[f(x_i) + \Delta x f'(x_i) + \frac{(\Delta x^2)}{2!} f''(x_i)\right]$, where a is a constant.

2 pts · Multiple choice · 5 alternatives

- (●) There is no error
- ( ) $\mathcal{O}(\Delta x)$.
- ( ) $\mathcal{O}(\Delta x^2)$.
- ( ) $\mathcal{O}(\Delta x^3)$.
- ( ) $\mathcal{O}(\Delta x^4)$.

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

Use Forward Euler to approximate the solution of the following Initial Value Problem from $t = 0$ until $t = 0.2$:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = y + \sin(t)$$
$$y(t_0 = 0) = 1$$
$$\Delta t = 0.1$$

6 pts · Open · 1 1/2 Page

$$\frac{y_{i+1} - y_i}{\Delta t} = y_i + \sin(t_i) \Rightarrow y_{i+1} = y_i + \Delta t \cdot (y_i + \sin(t_i))$$

Calculating for the first steps:

$$y_1 = 1 + 0.1 \cdot (1 + \sin(0)) = 1.1$$
$$y_2 = 1.1 + 0.1 \cdot (1.1 + \sin(0.1)) = 1.2199$$

6 pts

Correct expression of y(i+1)	2 pts
Correct expression of y(1) and y(2) including the correct time t(1) = 0.1 and t(2) = 0.2 for each timestep	2 pts
Correct final value of y(1)	1 pt
Correct final value of y(2)	1 pt

Is Backward Euler more accurate than Forward Euler?

2 pts · Multiple choice · 2 alternatives

- ○ Yes
- ◉ No

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

Consider the following convection equation:

$$\frac{dc}{\mathrm{d}t} + v\,\frac{dc}{dx}$$

Specify how many initial conditions and boundary conditions are required to solve the convection equation.

Then, write the algebraic expression to approximate the first derivative of this PDE using:

Central difference in space **and** central difference in time

5 pts · Open · 4/5 Page

It requires 1 initial condition and 1 boundary condition.

$$\frac{C_i^{j+1} - C_i^{j-1}}{2\Delta t} + v \cdot \frac{C_{i+1}^{j} - C_{i-1}^{j}}{2\Delta x}$$

5 pts

Correct amount of initial conditions	1 pt
Correct amount of boundary conditions	1 pt
Correct expression (and notation) of the Central difference in space **and** central difference in time used to approximate the first derivative of this PDE	3 pts

Consider a grid with 5 nodes: $x_0, x_{1,} x_{2,} x_3, x_4$

The temperature profile on this grid is described by the discretized ODE:

$$\frac{1}{\Delta x^2}\left(T_{i-1} - 2T_i + T_{i+1}\right) - \alpha\left(T_i - T_s\right) = 0$$

where α, Δx and T_s are known constants. The boundary conditions are Dirichlet:

$$T\left(x_0\right) = T_0 \text{ and } T\left(x_4\right) = T_4$$

A fellow student proposes the following system to solve the problem:

$$\begin{bmatrix} -(2+\alpha\Delta x^2) & 1 & 0 \\ 1 & -(2+\alpha\Delta x^2) & 1 \\ 0 & 1 & -(2+\alpha\Delta x^2) \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} \alpha T_s \Delta x^2 \\ \alpha T_s \Delta x^2 \\ \alpha T_s \Delta x^2 \end{bmatrix}$$

Explain what is wrong with this proposed system. How would you correct it?

6 pts · Open · 1 3/5 Page

The BCs are not implemented and there is a minus sign missing in the vector $\mathbf{b}$

$$\begin{bmatrix} -\alpha T_s \Delta x^2 - T_0 \\ -\alpha T_s \Delta x^2 \\ -\alpha T_s \Delta x^2 - T_4 \end{bmatrix}$$

6 pts

Adding the left BC correctly	2 pts
Adding the right BC correctly	2 pts
Adding the minus in the vector $\mathbf{b}$	2 pts
Correctly identify that BCs need to be added	1 pt
Minor error	-1 pts

Exercise 5: Probability and Reliability #43992455

25 pts · Last updated 24 Jan, 16:45

The empirical or sample mean based on m outcomes (or: realizations) x_i can be computed as:

$$\hat{\mu}_X = \frac{1}{m}\sum_{i=1}^{m} x_i$$

The sample variance can be computed as:

$$\hat{\sigma}_X^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \hat{\mu}_X)^2$$

Finally, the covariance is given by:

$$Cov(X_1, X_2) = \mathbb{E}\left(\left[X_1 - \mathbb{E}\left(X_1\right)\right]\left[X_2 - \mathbb{E}(X_2)\right]\right)$$

A group of scientists are analyzing the influence of ice water content in the clouds $\left[IC, kg/m^2\right]$ on rainfall $[RF, mm]$ as their theory is that it will improve the development of weather and climate models.

Here you have a brief definition of the variables:

- Ice water content in the clouds $\left[IC, kg/m^2\right]$. It is the amount of ice that you have in the clouds on the column of air above a given area. Thus, only positive values have a physical meaning.
- Rainfall $[RF, mm]$. Here, it is defined as the amount of rain that was accumulated during the hour when the measurement was conducted in the area used to compute IC. Again, only positive values have a physical
  meaning.

They have been able to conduct 10 field measurements that are shown in the table below:

Text

$IC\ [\mathrm{kg}/\ m^2]$	$RF\ [\mathrm{mm}]$
999	10
242	8
76	7
666	4
1292	12
927	9
405	9.5
126	7
867	6
491	10

Table 1: Pair observations of ice water content and rainfall.
Disclaimer: this example and data are generated for academic purposes.

Text

What do PDF and CDF stand for? Choose the right combination of words.

2 pts · Multiple choice · 4 alternatives

- ( ) P-value Distribution Function, and Cumulative Distribution Function
- (•) **Probability Density Function, and Cumulative Distribution Function**
- ( ) Probability Distribution Function, and Cumulative Density Function
- ( ) Probability Density Function, and Cumulative Density Function

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

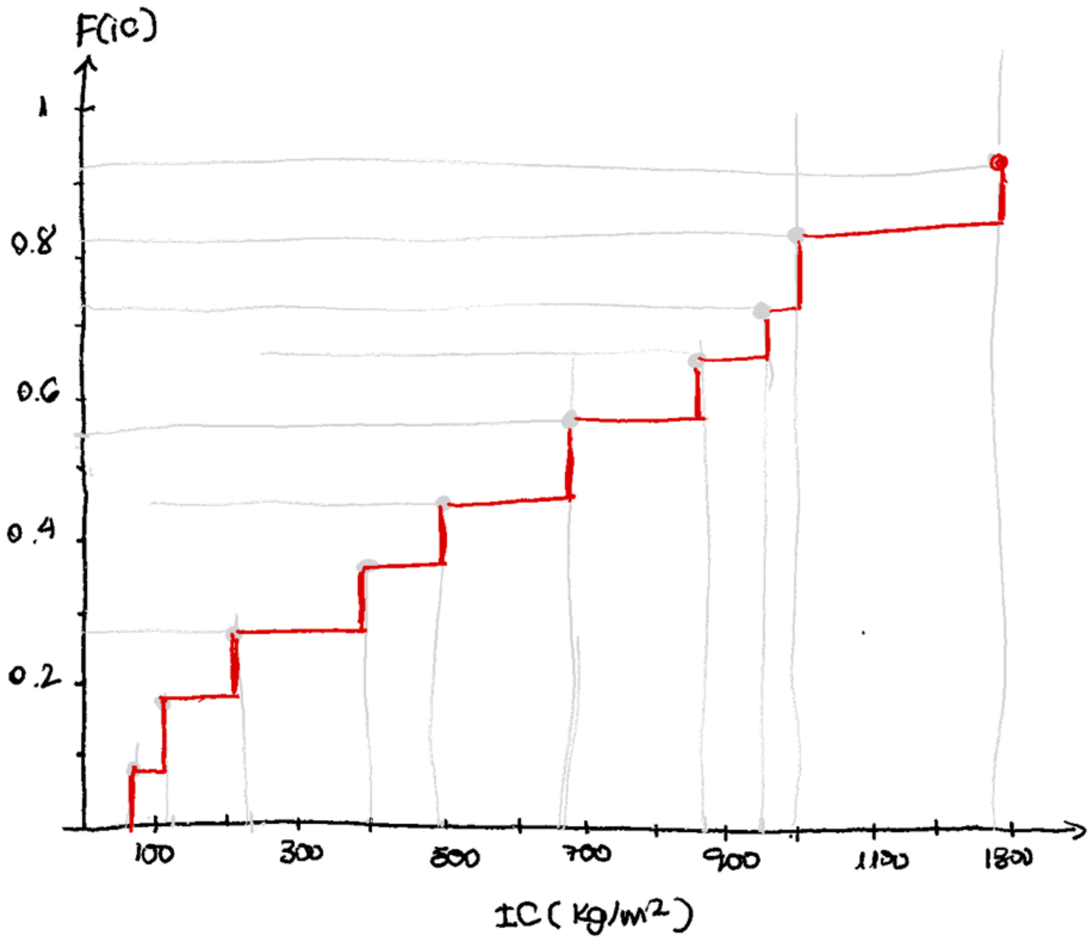
Compute and draw the empirical CDF of the variable IC.

Remember to label the axis and indicate their values. You may want to arrange your calculations in table format.

6 pts · Open · 1 7/10 Page

Model answer

IC (kg/m^2)	rank	$P[IC \leq ic]$
999	9	0.82
242	3	0.27
76	1	0.09
666	6	0.55
1292	10	0.91
927	8	0.73
405	4	0.36
126	2	0.18
867	7	0.64
491	5	0.45



Fully correct	6 pts
Assign ranks to the observations	2 pts
Incorrect or incomplete	-1 pts
No step function	-0.5 pts
Some indication of knowledge of what is a CDF: wrong plot but right axis, etc	1 pt
Other minor error	-0.5 pts
Correct table	4 pts

Is the Uniform distribution a reasonable distribution for the random variable IC? Justify your answer with at least one reason.

3 pts · Open · 3/5 Page

Model answer

Yes, it is. The above ECDF is approximately linear similar to a Uniform distribution. Also, we can compute the distance between different percentiles, such as $d(min, P55) = 666 - 76 = 590$ and $d(P55, max) = 1262 - 666 = 596$ and we can see that they are approximately the same so the empirical distribution seems to be symmetric. Also, the uniform distribution presents an upper and lower bound. In the case of the lower bound is desired although the upper bound does not represent reality.

Fully correct	3 pts
Reflection on (at least) one reason why Uniform distribution is suitable but not fully correct.	1.5 pts

Assuming that the scientists want to quantify the joint distribution function of IC and RF using a multivariate Gaussian distribution, what are the minimum parameters that they need to compute from the observations to quantify the distribution?

2 pts · Multiple choice · 4 alternatives

- ○ The mean, the variance and standard deviation of the each variable.
- ○ The mean of each variable and the correlation coefficient between them.
- ○ The covariance and correlation coefficient between the variables and the standard deviation of each variable.
- ◉ **The mean and variance of each variable and the covariance between them.**

Feedback

Feedback when the question is answered correctly

Feedback when the question is answered partially correctly

Feedback when the question is answered incorrectly

Assuming that the scientists want to quantify the joint distribution function of IC and RF using a multivariate Gaussian distribution.

Compute the covariance matrix of IC and RF. Do not use more than two decimal figures. You may want to make use of a table format to arrange your calculations.

8 pts · Open · 1 9/10 Page

Model answer

The covariance matrix of two variables, X and Y, is given by

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} 166,348.1 & 356 \\ 356 & 5.4 \end{pmatrix}$$

$$\hat{\mu}_{IC} = \frac{1}{10}(999 + 242 + 76 + 666 + 1292 + 927 + 405 + 126 + 867 + 491) = 609.1$$

$$\hat{\mu}_{RF} = 8.25$$

$$\hat{\sigma}^2_{IC} = \frac{1}{9}((999-609.1)^2 + (242-609.1)^2 + (76-609.1)^2 + (666-609.1)^2 + (1292-609.1)^2 + (927-609.1)^2 + (405-609.1)^2 + (126-609.1)^2 + (867-609.1)^2 + (491-609.1)^2) = 166,348.1$$

$$\hat{\sigma}^2_{RC} = 5.4$$

$$\mathrm{Cov}(X_1, X_2) = E\left[(X_1 - E(X_1))(X_2 - E(X_2))\right] = 356$$

IC (kg/m^2)	$IC_i - \mu_{IC}$	RF	$RF_i - \mu_{RF}$	$IC_i - \mu_{IC} \times RF_i - \mu_{RF}$
999	389.9	10	1.75	682.3
242	-367.1	8	-0.25	91.8
76	-533.1	7	-1.25	666.4
666	56.9	4	-4.25	-241.8
1292	682.9	12	3.75	2560.9
927	317.9	9	0.75	238.4
405	-204.1	9.5	1.25	-255.1
126	-483.1	7	-1.25	603.9
867	257.9	6	-2.25	-580.3
491	-118.1	10	1.75	-206.7
∑ / N				**356**

Fully correct	8 pts
Correct definition of covariance matrix (e.g.: non-zero covariance)	3 pts
Computation of variance correct	2 pts
Computation of covariance correct	3 pts
Computation error, minor error, computation not finished	-0.5 pts
Major error (e.g.: zero covariance, elements missing)	-1 pts

Assume that finally the scientists have changed their mind and do NOT want to quantify the joint distribution function of IC and RF using a multivariate Gaussian distribution. They have derived marginal distributions for IC, RF and the joint distribution of IC and RF as follows:

- The marginal distribution of IC is given by the PDF $f_{IC}\left(ic\right)$ and the CDF $F_{IC}\left(ic\right)$.
- The marginal distribution of RF is given by the PDF $f_{RF}\left(rf\right)$ and the CDF $F_{RF}\left(rf\right)$.
- The joint probability distribution of IC and RF is given by the PDF $f_{IC,RF}\left(ic, rf\right)$ and the CDF $F_{IC,RF}\left(ic, rf\right)$

Using the information described ONLY in this subquestion, derive the expression for the joint exceedance probability $P[IC > ic, RF > rf]$. You may want to draw a diagram to assist you.

4 pts · Open · 7/10 Page

Model answer

$$P[IC > ic, RF > rf] = 1 - F_{IC}(ic) - F_{RF}(rf) + F_{IC,RF}(ic, rf)$$

You have a zero if:

$$P[IC > ic, RF > rf] = 1 - P[IC < ic, RF < rf]$$

Fully correct	4 pts
Computed assuming independence between the random variables (although not indicated that they were) OR Error but good thinking behind	2 pts
Incomplete answer (e.g.: left in terms of probabilities)	3 pts
Minor error	-0.5 pts

extragrids.png

Image 1 Page

Use the grid below if you run out of space for any exercise.
In that case, please indicate so at the original answer field.