ST3009: Statistical Methods for Computer Science

Week 10 Assignment - Senán d'Art - 17329580

Q1

(a)

0
$F_X(x) - limy- > XF_X(y) = 0$
You cannot find the probability of continuous random variable having a discrete value as there are an uncountably infinite number of possible values.
In this case since we are analising a single point on the CDF curve, and there is no change at a point, the probability is 0.

(b)

Since $P(x < y) = y$ for $0 \leq x \leq 1$.
We can simply use the definition of the CDF.
The selected range is 0.25, which is a quarter of the total range and since the relationship is linear, we can determine:
$P(0.25 \leq x \leq 0.5) = P(0 \leq x \leq 0.25) = 0.25$

(c)

In this case we can disregard the case where $x < 0$. The new range becomes: $0 \leq x \leq 0.5$
Similarly to part (b) above:
$P(0 \leq x \leq 0.5) = 0.5$

Q2

(a)

We can picture the PDF as a line where the y axis goes from $0 \rightarrow 1$ and the x axis goes from $0 \rightarrow 2$.
The CDF can be viewed as the area beneath that line. We use the area of a triangle as $\frac{B}{2}H$. In this case $\frac{B}{2} = \frac{x}{2}$ and $H = \frac{x}{2}$.

$$CDF(x) = (\frac{x}{2})(\frac{x}{2}) = \frac{x^2}{4}$$

(b)

We can ignore the case where $x > 2$.

$0.5 \leq x \leq 2$

Using our formula from above (take $P(X < 0.5)$ from 1):

$1 - \frac{(0.5^2)}{4} = 0.9375$

Q3

(a)

Multiply the two probablities:

$$F_{XY}(x, y) = \frac{e^{-|x|}}{2} * e^{-2|y|}$$

(b)

$$F_{XY}(x, y) = \frac{e^{-|xy|}}{2}$$

From definition of conditional PDF:

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}$$

In our case:

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Where:

$$F_{XY}(x, y) = \frac{e^{-|xy|}}{2}$$

and

$$F_X(x) = \frac{e^{-|x|}}{2}$$

We get:

$$f_{Y|X}(y|x) = \frac{\frac{e^{-|xy|}}{2}}{\frac{e^{-|x|}}{2}} = \frac{e^{-|xy|}}{e^{-|x|}}$$

(c)

Definition of Bayes Rule for PDF:

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y) f_Y(y)}{f_X(x)}$$

In our case this becomes:

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}$$

Where:

$$f_{Y|X}(y|x) = \frac{e^{-|xy|}}{e^{-|x|}}$$

and:

$$f_Y(y) = e^{-2|y|}$$

and:

$$f_X(x) = \frac{e^{-|x|}}{2}$$

We get:

$$f_{X|Y}(x|y) = \frac{\frac{e^{-|xy|}}{e^{-|x|}} \frac{e^{-|x|}}{2}}{e^{-2|y|}} = \frac{\frac{e^{-|xy|}}{2}}{e^{-2|y|}} = \frac{e^{-|xy|}}{2e^{-2|y|}}$$

Q4

(a)

The joint PDFs are the product of all other PDFs ie:

$$f_{Z|X} = e^{-2|\theta y^{(1)} - x^{(1)}|} * e^{-2|\theta y^{(2)} - x^{(2)}|} * \dots * e^{-2|\theta y^{(m)} - x^{(m)}|}$$

Which becomes:

$$f_{Z|X} = \prod_{i=1}^{m} e^{-2|\theta y^{(i)} - x^{(i)}|}$$

(b)

1. Pick a random value for θ.
2. Test the value of θ in $f_{Z|X}$.
3. Test a value higher than θ by some value x and one lower than θ by x. This will allow us to see the slope and where the next estimate of θ should be.
4. Choose the value of θ which results in the largest $f_{Z|X}$ and repeat steps 2-4 until a desired accuracy is reached.

The value of x must be carefully chosen as a value too small will take a long time to reach an optimal answer and a value that is too large may miss the optimal answer completely.
It would be possible that the when $f_{Z|X}$ stops growing and starts shrinking we modify the algorithm. This could be done by taking the previous two values of θ and testing the midpoint. The

mid point then either becomes the upper or lower bound. We repeat this process until a desired accuracy is reached.