TRINITY COLLEGE DUBLIN

School of Computer Science and Statistics

Week 10 Questions ST3009: Statistical Methods for Computer Science

For each problem, explain/justify how you obtained your answer in order to obtain full credit. In fact, most of the credit for each problem will be given for the derivation/model used as opposed to the final answer.

Question 1. A continuous random variable X has CDF:

$$F_X(x) = \begin{cases} 1 & x > 1 \\ x & 0 \leq x \leq 1 \\ 0 & x < 0 \end{cases}$$

.

(a) Calculate the probability $P(X = 0.5)$ that X equals 0.5.
(b) Calculate the probability $P(0.25 \leq x \leq 0.5)$ that X lies between 0.25 and 0.5.
(c) Calculate the probability $P(-1 \leq x \leq 0.5)$.

Question 2. A continuous random variable X has PDF:

$$f_X(x) = \begin{cases} 0 & x > 2 \\ x/2 & 0 \leq x \leq 2 \\ 0 & x < 0 \end{cases}$$

.

(a) Calculate the CDF of X. Hint: Remember the area of a triangle is half the base times the height.
(b) Calculate $P(0.5 \leq X \leq 10)$.

Question 3. Continuous random variables X and Y have PDFs: $f_X(x) = \frac{e^{-|x|}}{2}$, $f_Y(y) = e^{-2|y|}$.

(a) Suppose X and Y are independent, write an expression for their joint PDF $f_{XY}(x, y)$.
(b) Suppose $f_{XY}(x, y) = \frac{e^{-|xy|}}{2}$ (so they're not independent), write an expression for the conditional PDF $f_{Y|X}(y|x)$. Hint: recall the definition of conditional PDF.
(c) Using Bayes Rule for PDFs and your answer from (b) write an expression for the conditional PDF $f_{X|Y}(x|y)$.

Question 4. Suppose you have a sequence of m random variables $Z = \{Y^{(i)}, i = 1, 2, \ldots, m\}$ with conditional PDF $f_{Y^{(i)}|X^{(i)}}(y^{(i)}|x^{(i)}) = e^{-2|\theta y^{(i)} - x^{(i)}|}$, where θ is an unknown parameter. The random variables $Y^{(i)}$, $i = 1, 2, \ldots, m$, are independent.

(a) Write an expression for the PDF $f_{Z|X}$ i.e. the joint PDF of the $Y^{(i)}$'s conditioned on the $x^{(i)}$'s, $i = 1, 2, \ldots, m$.
(b) Suppose you are given training data consisting of a set of m pairs of data $(x^{(i)}, y^{(i)})$, $i = 1, 2, \ldots, m$, where $y^{(i)}$ is an observation of RV $Y^{(i)}$. The PDF of this data is given by your answer from part (a), but the PDF depends on unknown parameter θ. Briefly discuss how you could use gradient ascent to choose a value of θ which maxmises the PDF of the training data.