CS 771A: Intro to Machine Learning, IIT Kanpur				**Quiz II** (19 Oct 2022)
Name	JAYA GUPTA			**20 marks**
Roll No	200471	Dept.	CSE	Page **1** of **2**

Instructions:
1. This question paper contains 1 page (2 sides of paper). Please verify.
2. Write your name, roll number, department above in **block letters neatly with ink**.
3. Write your final answers neatly **with a blue/black pen**. Pencil marks may get smudged.
4. Don't overwrite/scratch answers especially in MCQ – such cases will get straight 0 marks.
5. Do not rush to fill in answers. You have enough time to solve this quiz.

Q1. Write **T** or **F** for True/False **in the box** and give justification below. **(4 x (1+2) = 12 marks)**

1	If $A, B \subset \mathbb{R}^2$ are convex sets, then the set $C = \frac{(A+B)}{2} \stackrel{\text{def}}{=} \left\{\frac{(a+b)}{2}, a \in A, b \in B\right\}$ is always convex. Give a brief proof if True else give counter example if False.	T

By rules of convexity, we know
i) if A, B ⊂ convex, then (A + B) ⊂ convex.
ii) if A ⊂ convex set, then (A/2) ⊂ convex set.
So if (A+B) ⊂ convex, then $\frac{(A+B)}{2}$ ⊂ convex set.
Hence set C is convex.

2	The solution to $\min_{x\in\mathbb{R}} \frac{\lambda}{2}\cdot x^2 + \sum_{i=1}^{n}\left(x - a^i\right)^2$ approaches the mean of the real numbers $a^1, a^2, \ldots, a^n \in \mathbb{R}$ as $\lambda \to \infty$. Justify by deriving the solution below.	F

$f(x) = \frac{\lambda}{2}x^2 + \sum_{i=1}^{n} (x - a^i)^2$

$\frac{df}{dx} = \lambda x + 2\sum_{i=1}^{n}(x - a^i)$, $\frac{df}{dx} = 0$, $\boxed{x = \frac{2\sum_{i=1}^{n} a_i}{(2n+\lambda)}}$

$\frac{d^2f}{dx^2} = (\lambda + 2n)$

as $\lambda > 0$, $n \geq 0$ $\frac{\partial^2 f}{\partial x^2} > 0$, minima.

as $\lambda \to \infty$, $x \to 0$

$\boxed{x = 0}$ Ans

$\lim_{\lambda\to\infty} 2\sum_{i=1}^{n} a_i / (2n+\lambda) \to 0$

3	For a doubly differentiable fn $f: \mathbb{R} \to \mathbb{R}$, if $f''(x_0) = 0$ at $x_0 \in \mathbb{R}$, then it must be that $f'(x_0) = 0$. Give brief justification if True else give counter example if False.	F

Statement is False.

$f(x) = x$.

$f'(x) = 1$

$f''(x) = 0$

Now $f''(x_0) = 0 \; \forall x_0 \in R$, but $f'(x_0) \neq 0$.

$V[X] = Var[X]$

4	Let X, Y be two random variables (may or may not be independent) such that $\text{Var}[X] > 0$ and $\text{Var}[Y] > 0$. Then it may happen that $\text{Var}[X+Y] = 0$. Give an example if your answer is True else give a proof that $\text{Var}[X+Y] > 0$ always.	T

Let X be a random variable with $V[X] = \sigma^2 > 0$

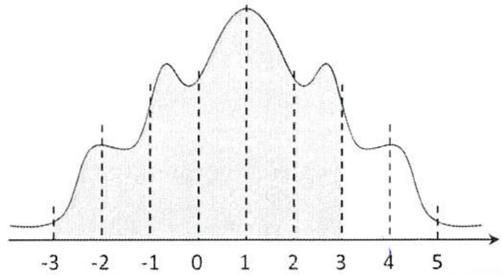
Take $Y = -X$. We know $V[cX] = c^2 V[X]$

$V[-X] = (-1)^2 V[X] = V[X] = \sigma^2 > 0$

So $V[X]$ and $V[Y]$ both are positive. But.

$V[X+Y] = V[X-X] = V[0] = \underline{0}$. Hence proved

Q2. (Deviant behaviour) Melbo created a new class of $\mathcal{M}$ distributions with mean μ, std σ such that if $X \sim \mathcal{M}(\mu, \sigma^2)$, then for any $c > 0$ we have $\mathbb{P}[\mu - c\sigma \leq X \leq \mu + c\sigma] = \eta_c$ It is known that $\eta_1 = 0.75, \eta_2 = 0.9, \eta_3 = 0.95, \eta_4 = 0.99$. For $\mu = 1, \sigma = 1$, find out $\mathbb{P}[-3 \leq X \leq 3]$. Give brief calculations below and the final answer. **(2+2=4 marks)**



$\mathbb{P}[1 - c \leq X \leq 1 + c] = \eta_c$, $\eta_1 = P[0 \leq X \leq 2] = 0.75$

$\eta_2 = P[-1 \leq X \leq 3] = 0.9$

$\eta_3 = P[-2 \leq X \leq 4] = 0.95$

$\eta_4 = P[-3 \leq X \leq 5] = 0.99$

$P[-3 \leq X \leq 3] = P[-1 \leq X \leq 3] + \underbrace{P[-3 \leq X \leq -1]}$

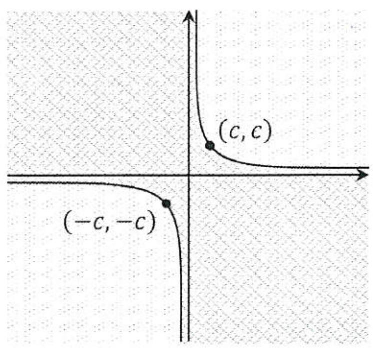
$P[-3 \leq X \leq -1] = \left(P[-3 \leq X \leq 5] - P[-1 \leq X \leq 3]\right)/2$

$P[-3 \leq X \leq -1] = \frac{0.99 - 0.9}{2} = 0.09/2 = 0.045$

$P[-3 \leq X \leq 3] = 0.9 + 0.045 = \boxed{0.945}$

Ans

Q3. (XOR classifier) For a 2D rectangular hyperbola with equation $xy = c^2$ for $c \in \mathbb{R}$, give a feature map $\phi: \mathbb{R}^2 \to \mathbb{R}^D$ for some $D > 0$ and a corresponding linear classifier $\mathbf{W} \in \mathbb{R}^D$ so that for any $\mathbf{x} \in \mathbb{R}^2$, $\text{sign}(\mathbf{W}^\top \phi(\mathbf{x}))$ takes value $z = -1$ in the light dotted region (see figure on the right) and $z = +1$ in the dark cross-hatched region. We don't care what happens on the boundary. Your map ϕ must **not depend on** c but $\mathbf{W}$ may depend on c. No need to show calculations. Just give the final answers. **(2 + 2 = 4 marks)**



$\phi(x, y) = \begin{bmatrix} 1 \\ xy \end{bmatrix}$

$\mathbf{W} = \begin{bmatrix} c^2 \\ -1 \end{bmatrix}$