Quiz 1	**Name:** ______________________
Week 2, Sep/17/2019	**On your left:** ______________________
CS 280: Fall 2019	**On your right:** ______________________
Instructor: Xuming He	

Instructions:
Please answer the questions below. Show all your work. This is an open-book test. NO discussion/collaboration is allowed.

Problem 1. (10 points) *Variance*. Assume $x_1, \cdots, x_n$ are independent random variables, show that

$$Var(x_1 + \cdots + x_n) = Var(x_1) + \cdots + Var(x_n)$$

Problem 2. (10 points) *Gradient*. Let $\sigma(a) = \frac{1}{1+e^{-a}}$ be an activation function, and $f(\mathbf{w}) = -\sum_{i=1}^{n}[y_i \log(\mu_i) + (1-y_i)\log(1-\mu_i)]$, where $\mu_i = \sigma(\mathbf{w}^\intercal \mathbf{x}_i)$.

(a) Show that $\frac{d}{d\mathbf{w}} f(\mathbf{w}) = \sum_{i=1}^{n}(\mu_i - y_i)\mathbf{x}_i$.

(b) (bonus 10 points) Show that $f(\mathbf{w})$ is convex.

Problem 3. (10 points) *Learning basics*. Consider Ridge regression

$$L(\mathbf{w}) = \sum_{i=1}^{n} \|\mathbf{w}^\intercal \mathbf{x}_i - y_i\|^2 + \lambda \mathbf{w}^\intercal \mathbf{w}$$

(a) Show that its optimal weight $\mathbf{w}^* = (X^\intercal X + \lambda I)^{-1} X^\intercal Y$.

(b) (bonus 10 points) Provide a probabilistic formulation of the objective function using the Bayesian Theorem.