CMPS 142 Machine Learning
Spring 2018, Homework #0

Aaron Steele, atsteele@ucsc.edu

Question 1

X is Normal, mean= μ, variance= σ^2.
The Probability Density Function (PDF) is

$$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Question 2

Probability for heads on a coin is λ.

A

The probability of getting a head is 1/2.
So, the probability for getting a head on the $(k+1)$-th toss is

$$\frac{1}{2}^{(k+1)}$$

B

The expected number of failures before first success is

$$E(X) = \frac{(1-p)}{p} = \frac{1-0.5}{0.5} = 1$$

Therefore, 2 tosses are expected to get the first head.

Question 3

A

Var(X) = $E[(X - E[X])^2]$. Prove Var(X) = $E[X^2] - E[X]^2$

$$E[(X - E[X])^2] = E[X^2 - 2XE[X] + E[X]^2]$$

$$= E[X^2] - 2E[XE[X]] + E[E[X]^2]$$
$$= E[X^2] - 2(E[X]E[X]) + E[X]^2$$
$$= E[X^2] - 2(E[X]^2) + E[X]^2$$
$$= E[X^2] - E[X]^2$$

B

Var(X) = $E[X^2] - E[X]^2$, so $1 - 0^2 = 1$.
Y = $a + bX$, Var(Y) = ?

$$Var(Y) = E[((a + bX) - E[a + bX])^2]$$
$$= E[((a + bX) - (a - b\mu)^2)]$$
$$= E[a^2(X - \mu)^2]$$
$$= a^2E[(X - \mu)^2]$$
$$Var(Y) = a^2Var(X), \text{ by prior problem.}$$

Question 4

A

$$\frac{\partial f}{\partial x} = 6x - y - 11$$
$$\frac{\partial f}{\partial y} = 2y - x$$

B

Find critical point of f(x,y).

$$6x - y - 11 = 0$$
$$2y - x = 0$$

Solve for y from the first equation

$$y = 6x - 11$$

Substitute y into the second equation

$$2(6x - 11) - x = 0 \Rightarrow 12x - 22 - x = 0 \Rightarrow 11x - 22 = 0 \Rightarrow x = 2$$

Substitute x back into the second equation

$$y = 6 \cdot 2 - 11 = 1$$

Now that we have the x,y pair, use the 2nd Derivative Test to determine what kind of critical point this is.

$$D = \frac{\partial^2 f}{\partial^2 x} = 6$$
$$z = 6(2,1) = 6$$

D > 0, and z > 0, which means it is a relative minimum. It also happens to be a global minimum.

Question 5

A

It is known that if the second derivative of f(x) is ≥ 0 then it is convex.

$$f'(x) = 2x,\ f''(x) = 2$$

Which is ≥ 0, therefore x^2 is convex.

B

$$x^T(\lambda A + (1-\lambda)B)x = \lambda x^T Ax + (1-\lambda)x^T Bx \geq 0$$
$$x^T(\lambda Ax + (1-\lambda)Bx) \geq 0$$
$$x^T x(\lambda A + (1-\lambda)B) \geq 0$$
$$\lambda A + (1-\lambda)B \geq 0$$

So, $x^T Ax$ is convex.

Question 6

Machine learning is a hot topic in the computing world, and I know pretty much nothing about it. When one of my friends asked me what the difference between ML and AI was I wasn't able to tell them. So, I want to know the difference.

I also want to be able to use the basics of it and more easily learn about it since it seems such a versatile tool.