CSE343/ ECE363/ ECE563: Machine Learning W2021 // **Assignment-0 Basic Tools for ML**

Navneet Agarwal
2018348

February 7, 2021

3 Probability Theory

1)

P(App developed by A) = $\frac{15}{100}$ P(App developed by B) = $\frac{20}{100}$

P(App developed by C) = $\frac{30}{100}$ P(App developed by D) = $\frac{35}{100}$

P(App developed by A had a bug) = $\frac{15}{100} \cdot \frac{8}{100}$ P(App developed by B had a bug) = $\frac{20}{100} \cdot \frac{5}{100}$

P(App developed by C had a bug) = $\frac{30}{100} \cdot \frac{4}{100}$ P(App developed by D had a bug) = $\frac{35}{100} \cdot \frac{2}{100}$

(a)

P(An app developed by the company is chosen randomly and is found to have bug was developed by A) =
. = P(App had a bug and it was developed by A)/P(App had a bug)

$$RequiredProbability = \frac{\frac{15}{100} \cdot \frac{8}{100}}{\frac{15}{100} \cdot \frac{8}{100} + \frac{20}{100} \cdot \frac{5}{100} + \frac{30}{100} \cdot \frac{4}{100} + \frac{35}{100} \cdot \frac{2}{100}}$$

$$RequiredProbability = \frac{120}{120 + 100 + 120 + 70}$$

$$RequiredProbability = \frac{12}{41}$$

(b)

P(An app developed by the company is chosen randomly and is found to have bug was developed by B) =
. = P(App had a bug and it was developed by B)/P(App had a bug)

$$RequiredProbability = \frac{\frac{20}{100} \cdot \frac{5}{100}}{\frac{15}{100} \cdot \frac{8}{100} + \frac{20}{100} \cdot \frac{5}{100} + \frac{30}{100} \cdot \frac{4}{100} + \frac{35}{100} \cdot \frac{2}{100}}$$

$$RequiredProbability = \frac{100}{120 + 100 + 120 + 70}$$

$$RequiredProbability = \frac{10}{41}$$

(c)

P(An app developed by the company is chosen randomly and is found to have bug was developed by C) =
. = P(App had a bug and it was developed by C)/P(App had a bug)

$$RequiredProbability = \frac{\frac{30}{100}\cdot\frac{4}{100}}{\frac{15}{100}\cdot\frac{8}{100}+\frac{20}{100}\cdot\frac{5}{100}+\frac{30}{100}\cdot\frac{4}{100}+\frac{35}{100}\cdot\frac{2}{100}}$$
$$RequiredProbability = \frac{120}{120+100+120+70}$$
$$RequiredProbability = \frac{12}{41}$$

(d)

P(An app developed by the company is chosen randomly and is found to have bug was developed by D) =
. = P(App had a bug and it was developed by D)/P(App had a bug)

$$RequiredProbability = \frac{\frac{35}{100}\cdot\frac{2}{100}}{\frac{15}{100}\cdot\frac{8}{100}+\frac{20}{100}\cdot\frac{5}{100}+\frac{30}{100}\cdot\frac{4}{100}+\frac{35}{100}\cdot\frac{2}{100}}$$
$$RequiredProbability = \frac{70}{120+100+120+70}$$
$$RequiredProbability = \frac{7}{41}$$

2)

Probability of getting a head = 2 times the probability of getting a tail
P(Head) = 2* P(Tail) and P(Head) + P(Tail) = 1
Thus, P(Head) = 2/3 and P(Tail) = 1/3

(a)

Probability that A wins = Probability of getting a head
. = 2/3

(b)

Expected value of points earned by A = $1 * P(Head) + 0 * P(Tail)$
. = 2/3

(c)

Let x be the number of points earned, then we have

$$P(x) = \begin{cases} 2/3 & \text{if } x = 1 \\ 1/3 & \text{if } x = 0 \end{cases} \tag{1}$$

This is a **bernoulli distribution** with p = 2/3

(d)

For bernoulli distribution,

$$Variance = p(1-p) \tag{2}$$
$$Variance = \frac{2}{3} * \frac{1}{3} = \frac{2}{9} \tag{3}$$

3)

(a)

Points achieved by A = $\binom{10}{r}.(\frac{2}{3})^r.(\frac{1}{3})^{10-r}$

Thus, it follows a Binomial Distribution with n=10,p=2/3 and q=1/3

(b)

Expected value of score achieved by A = n.p (Since, its a binomial distribution)
. $= 10 * \frac{2}{3}$
. $= \frac{20}{3}$

(c)

Variance of score achieved by A = n.p.q (Since, its a binomial distribution)
. $= 10 * \frac{2}{3} * \frac{1}{3}$
. $= \frac{20}{9}$

(d)

Probability that A wins atleast 2 tosses = 1 - P(A loses all tosses) - P(A wins exactly one toss)
. $=1 - \binom{10}{0}(\frac{1}{3})^{10} - \binom{10}{1}(\frac{2}{3})^1(\frac{1}{3})^9$

. $=1 - \frac{1}{3^{10}} - \frac{10*2}{3^{10}}$

. $=1 - \frac{21}{3^{10}}$

. $=1 - \frac{7}{3^9}$

. $= \frac{19676}{19683}$

4)

(a)

Pr[B=1] = P(Number on dice is prime)*P(coin shows a head) + P(Number on dice is not a prime)*P(Number
. on second roll of dice is 1)

Pr[B=1] = $\frac{3}{6} * \frac{1}{2} + \frac{3}{6} * \frac{1}{6}$

Pr[B=1] = $\frac{1}{4} + \frac{1}{12}$

Pr[B=1] = $\frac{1}{3}$

(b)

Expected value of A = $\frac{1}{6} * 1 + \frac{1}{6} * 2 + \frac{1}{6} * 3 + \frac{1}{6} * 4 + \frac{1}{6} * 5 + \frac{1}{6} * 6$

Expected value of A = $\frac{21}{6}$

Expected value of A = $\frac{7}{2}$

(c)i

E[B | A is prime] = $\frac{1*\frac{1}{2}*\frac{1}{2}}{\frac{1}{2}}$= $\frac{1}{2}$

(c)ii

E[B | A is not prime] = $\frac{[1+2+3+4+5+6]*\frac{1}{2}*\frac{1}{6}}{\frac{1}{2}}$= $\frac{7}{2}$

(c)iii

$$E[A|B=1] = \frac{[2+3+5]*\frac{1}{6}*\frac{1}{2}+[1+4+6]*\frac{1}{6}*\frac{1}{6}}{\frac{1}{3}} \tag{4}$$

$$E[A|B=1] = \frac{\frac{10}{12}+\frac{11}{36}}{\frac{1}{3}} \tag{5}$$

$$E[A|B=1] = \frac{41}{12} \tag{6}$$

5

Assumption: we are given with mean and standard deviation
Gaussian(60,15) means Mean,$\mu = 60$ and Standard Deviation, $\sigma = 15$

(a)

Average value of T = Mean = 60

(b)

Standard Deviation in T = 15

(c)

$$P[T > 75] = 1 - P[T \leq 75] \tag{7}$$

$$P[T > 75] = 1 - \phi(\frac{75-\mu}{\sigma}) \tag{8}$$

$$P[T > 75] = 1 - \phi(\frac{75-60}{15}) \tag{9}$$

$$P[T > 75] = 1 - \phi(1) \tag{10}$$

$$P[T > 75] = 1 - 0.8413 \tag{11}$$

$$P[T > 75] = 0.1587 \tag{12}$$

(d)

$$P[T < 30] = \phi(\frac{30-\mu}{\sigma}) \tag{13}$$

$$P[T < 30] = \phi(\frac{30-\mu}{\sigma}) \tag{14}$$

$$P[T < 30] = \phi(\frac{30-60}{15}) \tag{15}$$

$$P[T < 30] = \phi(-2) \tag{16}$$

$$P[T < 30] = 0.0228 \tag{17}$$

(e)

$$P[45 \leq T \leq 75] = \phi(\frac{75-\mu}{\sigma}) - \phi(\frac{45-\mu}{\sigma}) \tag{18}$$

$$P[45 \leq T \leq 75] = \phi(\frac{75-60}{15}) - \phi(\frac{45-50}{15}) \tag{19}$$

$$P[45 \leq T \leq 75] = \phi(1) - \phi(-1) \tag{20}$$

$$P[45 \leq T \leq 75] = 0.8413 - 0.1587 \tag{21}$$

$$P[45 \leq T \leq 75] = 0.6826 \tag{22}$$

4 Vector Calculus: Gradients

1)

(a)

$$f(x,y) = -x^4 + 4(x^2 - y^2) + 20 \tag{23}$$

$$\nabla f(x,y) = \begin{bmatrix} \frac{\delta}{\delta x}(-x^4 + 4(x^2 - y^2) + 20) \\ \\ \frac{\delta}{\delta y}(-x^4 + 4(x^2 - y^2) + 20) \end{bmatrix} \tag{24}$$

$$\nabla f(x,y) = \begin{bmatrix} -4x^3 + 8x \\ -8y \end{bmatrix} \tag{25}$$

(b)

$$f(x,y,z) = -2x^3 + 5yz + z^4 \tag{26}$$

$$\nabla f(x,y,z) = \begin{bmatrix} \frac{\delta}{\delta x}(-2x^3 + 5yz + z^4) \\ \\ \frac{\delta}{\delta y}(-2x^3 + 5yz + z^4) \\ \\ \frac{\delta}{\delta z}(-2x^3 + 5yz + z^4)) \end{bmatrix} \tag{27}$$

$$\nabla f(x,y,z) = \begin{bmatrix} -6x^2 \\ 5z \\ 5y + 4z^3 \end{bmatrix} \tag{28}$$

2)

$$F(x) = b^T x \quad where \quad x, b \epsilon R^n \tag{29}$$

$$F(x) = [b_1x_1 + b_2x_2 + ...b_nx_n] \tag{30}$$

$$F(x) = [f(x)] \quad where \quad f(x) = b_1x_1 + b_2x_2 + ...b_nx_n \tag{31}$$

$$\nabla_x F(x) = \begin{bmatrix} \frac{\delta}{\delta x_1} f(x) \\ \frac{\delta}{\delta x_2} f(x) \\ . \\ . \\ . \\ \frac{\delta}{\delta x_n} f(x) \end{bmatrix} \tag{32}$$

$$\nabla_x F(x) = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{bmatrix} \tag{33}$$

$$\nabla_x F(x) = b \tag{34}$$

3)

$$F(x) = x^T Ax \quad where \quad x\epsilon R^n \;\; \& \;\; A\epsilon R^{nxn} \tag{35}$$

$$F(x) = \begin{bmatrix} \Sigma_{i=1}^{i=n} x_i a_{i1} & \Sigma_{i=1}^{i=n} x_i a_{i2} & . & . & . & \Sigma_{i=1}^{i=n} x_i a_{in} \end{bmatrix} x \tag{36}$$

$$F(x) = \begin{bmatrix} \Sigma_{i=1}^{i=n} \Sigma_{j=1}^{j=n} x_i a_{ij} x_j \end{bmatrix} \tag{37}$$

$$F(x) = \left[\sum_{i=1}^{n} (a_{ii} x_i^2 + \sum_{i \neq j} x_i a_{ij} x_j)\right] \tag{38}$$

$$\nabla F(x) = \begin{bmatrix} \frac{\delta}{\delta x_1} (\Sigma_{i=1}^{n} (a_{ii} x_i^2 + \Sigma_{i \neq j} x_i a_{ij} x_j)) \\ \frac{\delta}{\delta x_2} (\Sigma_{i=1}^{n} (a_{ii} x_i^2 + \Sigma_{i \neq j} x_i a_{ij} x_j)) \\ . \\ . \\ . \\ \frac{\delta}{\delta x_n} (\Sigma_{i=1}^{n} (a_{ii} x_i^2 + \Sigma_{i \neq j} x_i a_{ij} x_j)) \\ \end{bmatrix} \tag{39}$$

$$\nabla F(x) = \begin{bmatrix} 2a_{11} x_1^2 + \Sigma_{i \neq 1} (x_i a_{i1}) + \Sigma_{i \neq 1} (a_{1i} x_i) \\ 2a_{22} x_2^2 + \Sigma_{i \neq 2} (x_i a_{i2}) + \Sigma_{i \neq 2} (a_{2i} x_i) \\ . \\ . \\ . \\ 2a_{nn} x_n^2 + \Sigma_{i \neq n} x_i a_{in} + \Sigma_{i \neq n} a_{ni} x_i \\ \end{bmatrix} \tag{40}$$

$$\nabla F(x) = \begin{bmatrix} \Sigma_{i=1}^{n} (x_i a_{i1}) + \Sigma_{i=1}^{n} (a_{1i} x_i) \\ \Sigma_{i=1}^{n} (x_i a_{i2}) + \Sigma_{i=1}^{n} (a_{2i} x_i) \\ . \\ . \\ . \\ \Sigma_{i=1}^{n} (x_i a_{in}) + \Sigma_{i=1}^{n} (a_{1i} x_i) \\ \end{bmatrix} \tag{41}$$

$$\nabla F(x) = \begin{bmatrix} \Sigma_{i=1}^{n} (x_i a_{i1}) \\ \Sigma_{i=1}^{n} (x_i a_{i2}) \\ . \\ . \\ . \\ \Sigma_{i=1}^{n} (x_i a_{in}) \\ \end{bmatrix} + \begin{bmatrix} \Sigma_{i=1}^{n} (a_{1i} x_i) \\ \Sigma_{i=1}^{n} (a_{2i} x_i) \\ . \\ . \\ . \\ \Sigma_{i=1}^{n} (a_{ni} x_i) \\ \end{bmatrix} \tag{42}$$

$$\nabla F(x) = A^T x + Ax \tag{43}$$

4)

$$F(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} -x + y^2 \\ \sin x + 10y^2 \end{bmatrix} \tag{44}$$

$$Jacobian, J = \begin{bmatrix} \frac{\delta}{\delta x}(-x + y^2) & \frac{\delta}{\delta y}(-x + y^2) \\ \\ \frac{\delta}{\delta x}(\sin x + 10y^2) & \frac{\delta}{\delta y}(\sin x + 10y^2) \end{bmatrix} \tag{45}$$

$$Jacobian, J = \begin{bmatrix} -1 & 2y \\ \cos x & 20y \end{bmatrix} \tag{46}$$