CMPT 825 Homework 0, Fall 2020

Coulton Fraser (301405411)

September 15, 2020

1. Probability 10pt

(a)

	Red	Green
Box A	1/3	2/3
Box B	1/3	2/3
Box C	1/3	2/3

$$\begin{aligned} P(Red) &= \frac{\sum\limits_{P(Red)}}{\sum\limits_{P(Red)} + \sum\limits_{P(Green)}} \\ &= \frac{3(1/3)}{3(1/3)+3(2/3)} \\ &= \frac{1}{3} \end{aligned} \tag{1}$$

(b)

	Red	Green
Box A	5/15	0
Box B	4/15	0
Box C	6/15	0

$$\begin{aligned} P(Red|Box1) &= \frac{P(Red|BoxA)}{P(Red|BoxA)+P(Red|BoxB)+P(Red|BoxC)} \\ &= \frac{1}{3} \end{aligned} \tag{2}$$

2. Probability 15pt

(a)

$$\begin{aligned} P(Pos) &= P(TruePos) + P(FalseNeg) \\ &= (0.005 \cdot 0.98) + (0.995 \cdot 0.03) \\ &= 0.03475 \\ &= 3.475\% \end{aligned} \tag{3}$$

(b)

$$\begin{aligned} P(Disease|Pos) &= \frac{P(Pos|Disease) + P(Disease)}{P(Pos)} \\ &= \frac{0.98 \cdot 0.005}{0.03475} \\ &= 0.141007 \\ &\approx 14.101\% \end{aligned} \tag{4}$$

3. Probability 10pt

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_{1}^{2} x \cdot 2x^{-2}dx \\ &= 2\int_{1}^{2} \frac{1}{x}dx \\ &= 2\Big[\ln|x|\Big]_1^2 \\ &= 2\Big[\ln|2| - \ln|1|\Big] \\ &= 2ln|2| \\ &\approx 1.386 \end{aligned} \tag{5}$$

$$\begin{aligned} Var(X) &= E(X^2) - E(X)^2 \\ E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx \\ &= \int_{1}^{2} x^2 \cdot 2x^{-2}dx \\ &= 2\int_{1}^{2} x^0 dx \\ &= 2\Big[x\Big]_1^2 \\ &= 2\Big[2-1\Big] \\ &= 2 \\ Var(X) &= 2 - (2\ln(2))^2 \\ &\approx 0.078 \end{aligned} \tag{6}$$

4. Derivatives 10pt

(a)

$$f'(x) = 5(3x+1)^4 \tag{7}$$

(b)

$$f'(x) = \frac{8x}{4x^2+9} \tag{8}$$

(c)

$$
\begin{aligned}
f'(u) &= e^{-u} \cdot u' \\
f'(x) &= -e^{-x}
\end{aligned} \tag{9}
$$

(d)

$$
\begin{aligned}
d/dx\left[\frac{f}{g}\right] &= \frac{gf' - fg'}{g^2} \\
f'(x) &= \frac{4x^2 - 8x^2 - 2x}{x^4} \\
&= -\frac{4x+2}{x^3}
\end{aligned} \tag{10}
$$

(e)

$$
f'(x) = -\frac{1}{x^2} \tag{11}
$$

5. Matrix 10pt

(a)

$$
\begin{aligned}
1 \cdot 0 + -4 \cdot 5, 1 \cdot 3 + -4 \cdot 8 \\
-2 \cdot 0 + 1 \cdot 5, -2 \cdot 3 + 1 \cdot 8 \\
\begin{bmatrix} -20 & -29 \\ 5 & 2 \end{bmatrix}
\end{aligned} \tag{12}
$$

(b)

$$
\begin{aligned}
3 \cdot 1 + -2 \cdot -5 + 2 \cdot 8, 3 \cdot -3 + -2 \cdot 2 + 2 \cdot 3 \\
7 \cdot 1 + 1 \cdot -5 + -2 \cdot 4, 7 \cdot 3 + 1 \cdot 2 + -2 \cdot 3 \\
\begin{bmatrix} 29 & -11 \\ -6 & -25 \end{bmatrix}
\end{aligned} \tag{13}
$$

6. Matrix 10pt

(a)

$$
\begin{aligned}
&\vec{y} = W\big[C \times D\big] \cdot \vec{x}\big[D \times 1\big] \\
&\vec{y}\big[C \times 1\big]
\end{aligned} \tag{14}
$$

$\vec{y}$ Will have a length of C rows by one column.

(b)

$$
\begin{aligned}
\vec{y_i} &= W_{i,d}\vec{x}_d \\
\vec{y_3} &= W_{3,7}\vec{x}_7 \\
\frac{\partial_{y3}}{\partial_{x7}} &= W_{3,7}\partial_{x7}
\end{aligned} \tag{15}
$$

(c)

$$M' = \begin{bmatrix} 1 & 1 \\ 2x & e^x \\ 0 & 0 \end{bmatrix} \tag{16}$$

7. Information Theory 10pt

(a)

$$\begin{aligned} H(X) &= -\sum_{x \in X} p(x) log p(x) \\ H(P) &= -\Big[0.1 log(0.1) + 0.98 log(0.98) 0.1 log(0.1)\Big] \\ &\approx 0.6929 \\ H(Q) &= -\Big[0.15 log(0.15) + 0.60 log(0.60) 0.25 log(0.25)\Big] \\ &\approx 1.3527 \end{aligned} \tag{17}$$

(b)

$$\begin{aligned} H(P, Q) &= -\sum_{x \in X} p(P) log p(Q) \\ &= -\Big[0.1 log(0.15) + 0.98 log(0.60) 0.1 log(0.25)\Big] \\ &\approx 0.7696 \end{aligned} \tag{18}$$

(c) It is minimized when P=Q. It is maximized when P and Q have the least amount of mutual information or difference.

8. Information Theory 10pt

(a)

$$\begin{aligned} H(X) &= -\sum_{x \in X} p(x) log p(x) \\ &= -\Big[2(\frac{1}{8} log(\frac{1}{2})) + 6(\frac{1}{16} log(\frac{1}{4})) + 4(\frac{1}{32} log(\frac{1}{4})) + \frac{1}{4} log(\frac{1}{4})\Big] \\ &= 1.75 \\ H(X|Y) &= \sum_{x,y} p(x, y) log \frac{1}{p(x|y} \\ &= 1.375 \\ MI &= H(X) - H(X|Y) \\ &= 0.375 \end{aligned} \tag{19}$$

(b) $0 \geq MI \leq \min(H(X), H(Y))$

9. Cosine Similarity 10pt

$$\begin{aligned}\cos\theta &= \frac{2\times 1 + 2\times 3}{\sqrt{2^2+2^2}\times\sqrt{1^2+3^2}} \\ &= 0.8944\end{aligned} \tag{20}$$