Q A 1.

A 5+14+5=24

5/5

i T since given $Y=1$, x either equals to 0 or 1.

ii F consider $\begin{array}{cc} & Y \\ X & \begin{matrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{matrix}\end{array}$ then $p(x=1|Y=1)+p(x=1|Y=0)$

$= 0.8+0.3 \neq 1$

iii F since $p(X=1,Y=1,Z=0) = 1-p(X=1,Y=1,Z=1)$

$= 1-p(X=1)p(Y=1)p(Z=1)$

$= 1-p(X=1)p(Y=1)(1-p(Z=0))$

$= 1-p(X=1)p(Y=1)+p(X=1)p(Y=1)p(Z=0)$

$= p(X=1)p(Y=1)p(Z=0)$ iff

$1-p(X=1)p(Y=1)=0$ which is not always true

iv. T $\cancel{\frac{p(X=0|Y=0)}{p(X=0,Y=1)}}$ $\dfrac{p(X=0,Y=0)}{p(X=0,Y=1)} = \dfrac{p(Y=0,X=0)}{p(Y=1,X=0)}$

$= \dfrac{p(Y=0|X=0)\,p(X=0)}{p(Y=1|X=0)\,p(X=0)}$

$= \dfrac{p(Y=0|X=0)}{p(Y=1|X=0)}$

v. T $p(X=0,Y=0)+p(X=1,Y=0) = \cancel{\dfrac{p(X=0,Y=0)\,p(X=0,Y=0)+p(X=1,Y=0)\,p(X=0,Y=0)}{p(X=0,Y=0)}}$

$\cancel{= \dfrac{p(x=0,Y=0)\,p(}{p(X=0|Y=0)\,p(Y=0)}}$

$= p(Y=0)$

$= \dfrac{p(X=0,Y=0)}{p(X=0,Y=0)}\,p(Y=0)$

$= \dfrac{p(X=0,Y=0)\,\cancel{p(Y=0)}}{p(X=0|Y=0)\,\cancel{p(Y=0)}}$

$= \dfrac{p(X=0,Y=0)}{p(X=0|Y=0)}$

Q A2.

14/14

a $p(w=1) = 0.9$

$p(w=1 \mid h=1) = 0.95$

$p(w=1 \mid h=0) = 0.85$.

i ~~$p(h=1) =$~~

we know $p(w=1) = p(w=1 \mid h=1)\, p(h=1) + p(w=1 \mid h=0)\, p(h=0)$

so $p(w=1) = p(w=1 \mid h=1)\, p(h=1) + p(w=1 \mid h=0) - p(w=1 \mid h=0)\, p(h=1)$

sub. numbers in $0.9 = 0.95\, p(h=1) + 0.85 - 0.85\, p(h=1)$

$0.05 = 0.1\, p(h=1)$

so $p(h=1) = 0.5$

5

ii $p(h=1 \mid w=1) = \dfrac{p(w=1 \mid h=1)\, p(h=1)}{p(w=1)}$ using Bayes theorem

$= \dfrac{0.95 \times 0.5}{0.9}$

$= \dfrac{19}{36}$

3

iii $p(I=1 \mid h=1) = 0.9$

$p(w=1 \mid h=1, I=1) = 2\, p(w=1 \mid h=1, I=0)$

we want

$p(w=1 \mid h=1, I=0)$

6

we know that ~~$p(W=1) = p$~~

$p(w=1 \mid h=1) = p(w=1, I=1 \mid h=1)$ ~~$+ p(w=1, I=0$~~ $+ p(w=1, I=0 \mid h=1)$

$= p(w=1 \mid I=1, h=1)\, p(I=1 \mid h=1) + p(w=1 \mid I=0, h=1)\, p(I=0 \mid h=1)$

$= 2\, p(w=1 \mid h=1, I=0)\, p(I=1 \mid h=1) + p(w=1 \mid I=0, h=1)(1 - p(I=1 \mid h=1))$

substitute numbers in:

$0.95 = 2 \times 0.9 \times p(w=1 \mid h=1, I=0) + 0.1\, p(w=1 \mid I=0, h=1)$

we get $p(w=1 \mid h=1, I=0) = \dfrac{0.95}{1.9} = \dfrac{1}{2}$

Q A3

5/6

i. the maximum likelihood estimate of θ_h is 0.6 ✓ 2

ii. As the number of trials N increase, the likelihood estimates approaches true probabilities of the underlying distribution. 1

LLN?

iii.
$$p(\theta|D') = \frac{p(D'|\theta)p(\theta)}{p(D')}$$ Bayes theorem 2

which expresses the posterior probability in terms of the prior and the evidence, doesn't depend on D' ~~alone~~ alone.

max $p(\theta|D')$

Q B 1

B 12+8+4=24

12/12

i) $H(X) = \frac{1}{2}\log_2 2 + \frac{1}{3}\log_2 3 + 2\times\frac{1}{12}\log_2 12$
$= \frac{1}{2} + \frac{1}{3}\log_2 3 + \frac{1}{6}\log_2 4 + \frac{1}{6}\log_2 3$
$\approx \frac{1}{2} + \frac{1}{3} + \frac{1}{2}\times 1.58$
$= 1.623$

2

ii) ~~p(Y) = (~~
$p(Y = \text{Naks}) = \frac{1}{6}\times\frac{1}{2} + \frac{1}{4}\times\frac{1}{3} + \frac{1}{2}\times\frac{1}{12}\times 2$
$= \frac{3}{12} = \frac{1}{4}$
$p(Y = \text{Ray}) = 1 - \frac{1}{4} = \frac{3}{4}$
Hence $H(Y) = \frac{1}{4}\log_2 4 + \frac{3}{4}\log_2\frac{4}{3}$
$= \frac{1}{2} + \frac{3}{2} - \frac{3}{4}\log_2 3$
$\approx 2 - \frac{3}{4}\times 1.58$
≈ 0.8

3

iii $H(Y \mid X = a) \approx 0.65 < 0.8 < 1 = H(Y \mid X = d)$

Note $H(Y \mid X = d) = \log_2 2 = 1$
and $H(Y \mid X = a) = \frac{1}{6}\log_2 6 + \frac{5}{6}\log_2\frac{6}{5}$
$= \frac{1}{6} + \frac{1}{6}\log_2 3 + \frac{5}{6}\log_2 6 - \frac{5}{6}\log_2 5$
$= \frac{1}{6} + \frac{5}{6} + \log_2 3 - \frac{5}{6}\log_2 5$
$\approx 1 + 1.58 - \frac{5}{6}\times 2.32$
~~4.54~~
≈ 0.65

3

iv. because [H(Y|X) contains less information] than H(Y) since given X we are less uncertain about Y (unless $X \perp\!\!\!\perp Y$, for which $H(Y \mid X) = H(Y)$)

?

2

v. $H(Z \mid X) = 0$ since ~~$p(Z \mid X) = (1, 0)$ and $p(Z \mid X = d) =$~~ $p(Z \mid X = a, b, c) = (1, 0)$ and $p(Z \mid X = d) = (0, 1)$
this is because z is deterministic in terms of X. given X there are no uncertainty left for the value of z.

2

Q B2.

8/9

i if (X, Y, Z) form a Markov chain, then

$p(X, Y, Z) = p(X)p(Y|X)p(Z|Y)$ and $X \perp\!\!\!\perp Z \mid Y$

but then

$p(Z, Y, X) = p(Z)p(Y|Z)p(X|Y, Z)$

$= p(Z)p(Y|Z)p(X|Y)$ since $X \perp\!\!\!\perp Z \mid Y$

hence (Z, Y, X) also form a Markov chain

5

ii the data processing inequality states that (for the markov chain (X, Y, Z))

$I(X;Y) \geq I(X;Z)$

since (Z, Y, X) also forms a Markov chain, ~~and~~ we have

$I(Z;Y) \geq I(X;Z)$

and so we get $I(X;Z) \leq \min(I(X;Y), I(Y;Z))$

since mutual information is symmetric

intuition?

2

iii let $X, Y, Z = \{0, 1\}$ ~~and $y = \{0$ if $x = 0, 1$ with 1~~

and if $x=0$, then $y=0$ or 1 with equal chance,

same when $x=1$. $p(X) = (0.5, 0.5)$

and for $y=0$, $z=0$ and $y=1$ implies $z=1$.

then

2

$I(X;Y) = H(Y) - H(Y|X)$

$= 1 - 1 = 0$

and $I(Y;Z) = H(Z) - H(Z|Y)$

$= 1 - 0 = 1$

i.e. $I(X;Y) < I(Y;Z)$

Q B3.

4/4

i $E[Z] = E[X] + E[Y] = 8000$

ii $p(Z \geq a) < \frac{E[Z]}{a}$

$p(Z \geq 20{,}000) < \frac{E[Z]}{20{,}000}$

$= 0.4$

2

iii since X and Y are dependent, exploring every possible combination and count is tedious.