Kyle McGraw

Bi 195 Set 2

$X = Y = Z = \{0,1\} \qquad P(X=1) = P(Y=1) = p$

1. $P(X,Y,Z) = P(X)P(Y)P(Z) \quad P(Z=1) = p$

a) $E[X] = \sum_{x \in X} x \cdot P(x) = 1 \cdot p + 0 \cdot (1-p) = p$

b)
$$
\begin{aligned}
\mathrm{Cov}(X,Z) &= E[(X - E[X])(Z - E[Z])] \\
&= E[(X-p)(Z-p)] \\
&= E[XZ] - E[Xp] - E[Zp] + E[p^2] \\
&= E[X]E[Z] - pE[X] - pE[Z] + p^2 \\
&= p^2 - p^2 - p^2 + p^2 \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(X,Y) &= E[(X - E[X])(Y - E[Y])] \\
&= E[(X-p)(Z-p)] \\
&= E[XY] - E[Xp] - E[Yp] + E[p^2] \\
&= E[X]E[Y] - pE[X] - pE[Y] + p^2 \\
&= p^2 - p^2 - p^2 + p^2 \\
&= 0
\end{aligned}
$$

c)
$$
\begin{aligned}
I[X,Y] &= \sum_{y \in Y} \sum_{x \in X} P(x,y) \log \frac{P(x,y)}{P(x)P(y)} \\
&= \sum_{y \in Y} \sum_{x \in X} P(x)P(y) \log \frac{P(x)P(y)}{P(x)P(y)} \\
&= \sum_{y \in Y} \sum_{x \in X} P(x)P(y) \log 1 \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
I[X,Z] &= \sum_{z \in Z} \sum_{x \in X} P(x,z) \log \frac{P(x,z)}{P(x)P(z)} \\
&= \sum_{z \in Z} \sum_{x \in X} P(x)P(z) \log \frac{P(x)P(z)}{P(x)P(z)} \\
&= \sum_{z \in Z} \sum_{x \in X} P(x)P(z) \log 1 \\
&= 0
\end{aligned}
$$

2. $Z = OR(X,Y)$

a)

$P(Z,X) = P(X)\,P(Z|X)$

X	Z	$P(X)$	$P(Z\|X)$	$P(Z,X)$	$P(Z)$	$P(X)P(Z)$
0	0	$1-p$	$1-P(Y) = 1-p$	$(1-p)^2$	$(1-p)^2$	$(1-p)^3$
0	1	$1-p$	$P(Y) = p$	$p(1-p)$	$1-(1-p)^3$	$(1-p)-(1-p)^3$
1	0	p	0	0	$(1-p)^2$	$p(1-p)^2$
1	1	p	1	p	$1-(1-p)^2$	$p - p(1-p)^2$
			1	sum $= (1-p)^2 + p(1-p) + p$		

$= (1-p)(1-p+p)+p$
$= 1-p+p$
$= 1$ ✓

$P(z,x) \neq P(x)P(z)$

b)

x	z	$P(x)$	$P(x \mid z)$	$P(z \mid x)$
0	0	$1-p$	1	$1-p$
0	1	$1-p$	$\frac{1-p}{2-p}$	p
1	0	p	0	0
1	1	p	$\frac{1}{2-p}$	1

$$P(x \mid z) = \frac{P(z,x)}{P(z)}$$

$$\frac{p(1-p)}{1-(1-p)^2} = \frac{p(1-p)}{2p-p^2} = \frac{1-p}{2-p}$$

$$\frac{p}{1-(1-p)^2} = \frac{p}{2p-p^2} = \frac{1}{2-p}$$

c) $$I[x,z] = \sum_{z\in Z}\sum_{x\in X} P(x,z)\log\frac{P(x,z)}{P(x)P(z)}$$

$$= (1-p)^2\log\frac{(1-p)^2}{(1-p)^3} + p(1-p)\log\frac{p(1-p)}{(1-p)-(1-p)^3} + p\log\frac{p}{p-p(1-p)^2}$$

$$= (1-p)^2\log\frac{1}{1-p} + p(1-p)\log\frac{p}{1-(1-p)^2} + p\log\frac{1}{1-(1-p)^2}$$

$$= (1-p)^2\log\frac{1}{1-p} + p(1-p)\log\frac{1}{2-p} + p\log\frac{1}{2p-p^2}$$

$$I[x,y] = \sum_{y\in Y}\sum_{x\in X} P(x,y)\log\frac{P(x,y)}{P(x)P(y)}$$

$$= \sum_{y\in Y}\sum_{x\in X} P(x)P(y)\log\frac{P(x)P(y)}{P(x)P(y)}$$

$$= \sum_{y\in Y}\sum_{x\in X} P(x)P(y)\log 1$$

$$= 0$$

d) $$\mathrm{Cov}(X,Z) = E[(X-E[X])(Z-E[Z])]$$

$E[x] = p$
$E[z] = 2p-p^2$

$$= E[XZ] - E[X]E[Z]$$

$$= \sum_{x\in X}\sum_{z\in Z} xz\,p(x,z) - 2p^2 + p^3$$

$$= p - 2p^2 + p^3$$

$$= p(1-2p+p^2)$$
$$= p(1-p)^2$$

3.

a) $Z = \mathrm{AND}(X,Y)$

X	Z	$P(x)$	$P(z\mid x)$	$P(z,x)$	$P(z)$	$P(x)P(z)$
0	0	$1-p$	1	$1-p$	$1-p^2$	$(1-p)(1-p^2)$
0	1	$1-p$	0	0	p^2	$(1-p)p^2$
1	0	p	$1-p$	$p(1-p)$	$1-p^2$	$p(1-p^2)$
1	1	p	p	p^2	p^2	p^3

$$I[X,Z] = \sum_{z\in Z}\sum_{x\in X} P(x,z)\log\frac{P(x,z)}{P(x)P(z)}$$

$$= (1-p)\log\frac{1-p}{(1-p)(1-p^2)} + p(1-p)\log\frac{p(1-p)}{p(1-p^2)} + p^2\log\frac{p^2}{p^3}$$

$$= (1-p)\log\frac{1}{1-p^2} + p(1-p)\log\frac{1}{1+p} + p^2\log\frac{1}{p}$$

$$Cov(X,Z) = E[(X-E[X])(Z-E[Z])]$$

$E[x]=p$
$E[z]=p^2$

$$= E[XZ]-E[X]E[Z]$$

$$= \sum_{x\in X}\sum_{z\in Z} xz\,p(x,z) - p^3$$

$$= p^2-p^3$$

$$= p^2(1-p)$$

b) $Z = XOR(X,Y)$

X	Z	$P(x)$	$P(z\mid x)$	$P(z,x)$	$P(z)$	$P(x)P(z)$
0	0	$1-p$	$1-p$	$(1-p)^2$	$p^2+(1-p)^2$	$(1-p)(p^2+(1-p)^2)$
0	1	$1-p$	p	$p(1-p)$	$2p(1-p)$	$2p(1-p)^2$
1	0	p	p	p^2	$p^2+(1-p)^2$	$p^3+p(1-p)^2$
1	1	p	$1-p$	$p(1-p)$	$2p(1-p)$	$2p^2(1-p)$

$1-2p+2p^2$
$2p-2p^2$

$$I[X,Z] = \sum_{z\in Z}\sum_{x\in X} P(x,z)\log\frac{P(x,z)}{P(x)P(z)}$$

$$= (1-p)^2\log\frac{(1-p)^2}{(1-p)(p^2+(1-p)^2)} + p(1-p)\log\frac{p(1-p)}{2p(1-p)^2} + p^2\log\frac{p^2}{p^3+p(1-p)^2}$$

$$+ p(1-p)\log\frac{p(1-p)}{2p^2(1-p)}$$

$$= (1-p)^2\log\frac{1-p}{p^2+(1-p)^2} + p(1-p)\log\frac{1}{2(1-p)} + p^2\log\frac{p}{p^2+(1-p)^2}$$

$$+ p(1-p)\log\frac{1}{2p}$$

$$Cov(X,Z) = E[(X-E[X])(Z-E[Z])]$$

$E[X] = p$

$E[Z] = 2p(1-p)$

$$= E[XZ] - E[X]E[Z]$$

$$= \sum_{x \in X} \sum_{z \in Z} xz\,p(x,z) - 2p^2(1-p)$$

$$= p(1-p) - 2p^2(1-p)$$

$$= (p - 2p^2)(1-p)$$

$$= p(1-p)(1-2p)$$

4.

In a biological system, mutual information would provide the most information about X, Y, Z because it will uncover non-linear correlations. Pitfalls could be that mutual information needs to know the probability distribution, which may be unknown in the biological system.