STAT 241, HOMEWORK #1, 9/19/2011

Mike Schachter
mike.schachter@gmail.com

Problem 1

Let $X_C(n), X_L(n)$ be the events representing the # of heads Consuela or Larry have by trial n, respectively, so that the event where Consuela flips more heads than Larry on her n+1 toss is $E=\{X_C(n+1)>X_L(n)\}$. There are three possible events that could exist after n coin flips:

$$A=\{X_C(n)=X_L(n)\}, B=\{X_C(n)<X_L(n)\}, C=\{X_C(n)>X_L(n)\}$$

Let $P(A)=p, P(B)=q, P(C)=q$. Since these events comprise the entire event space, $P(A)+P(B)+P(C)=1$, and $2\text{p}+q=1$.

We can condition on the past n trials and use the law of total probability to compute $P(E)$:

$$P(E)=P(E|A)P(A)+P(E|B)P(B)+P(E|C)P(C)$$

If Consuela is behind Larry (event B), she can't make it up in one coin flip so $P(E|B)=0$. If Consuela is already ahead (event C), so that $P(E|C)=1$. If the # of heads is equal on toss n, (event A), then $P(E|A)=P(Consuela\,flips\,heads\,on\,n+1)=\frac{1}{2}$. Combining these equations gives:

$$P(E)=(\frac{1}{2})p+q=(\frac{1}{2})(2\text{p}+q)=(\frac{1}{2})$$.

Reference: answer adapted from "The Cuny Math Contest" (http://math.cisdd.org/samples.php)

Problem 2

(a) Let $X_t \in \{1,2,3,4,5\}$ be the stochastic process that represents the door John takes at time t. The joint probability of the history of doors taken at time t is given as:

$$P(X_{1,..},..,X_t)=\prod_{k=1}^{t} P(X_k|X_{k-1},...,X_1)=\prod_{k=1}^{t} P(X_k|X_{k-1})$$

Because the joint probability of the process can be factored like this, X_t is a Markov chain.

(b) Assume the doors are arranged in a circle, so each one has a left and right neighbor; door 5 is the left neighbor of 1. The transition probability matrix, where $T_{ij}=P(X_{k+1}=j|X_k=i)$ is given as:

$$\boldsymbol{T}=\begin{bmatrix} 0 & p & 0 & 0 & (1-p) \\ 1-p & 0 & p & 0 & 0 \\ 0 & 1-p & 0 & p & 0 \\ 0 & 0 & 1-p & 0 & p \\ p & 0 & 0 & 1-p & 0 \end{bmatrix}$$

(c) The steady state probability is computed as $T^{\infty} = \lim_{n\to\infty} \prod_{1}^{\infty} T$. Maybe there's a smart way of analytically computing the matrix power, like diagonalizing it first by finding the eigenvectors... but 5x5 matricies are kind of frustrating to deal with, so I'll provide a numerical result. A small R program was written to approximate the steady state transition matrix:

```
steady_state = function(n, p)
{
  q = 1 - p
  T = array(0, c(5, 5))
  T[1, 2] = p; T[2, 3] = p; T[3, 4] = p; T[4, 5] = p; T[5, 1] = p;
  T[1, 5] = q; T[2, 1] = q; T[3, 2] = q; T[4, 3] = q; T[5, 4] = q  ;
  Tn = T
  for (k in 1:n) { Tn = Tn%*%T }
  print(T)
  print(Tn)
}
```

This program gives the following matrix for $n=100, p\in(0,1)$:

$$\boldsymbol{T} = \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}$$

Problem 3

(a) The statement to prove/disprove is:

$$X_1 \perp X_4 | \{X_2, X_3\} \quad \text{implies} \quad X_2 \perp X_3 | \{X_1, X_4\}$$

This statement is false, a counterexample is given by the following graphical model:

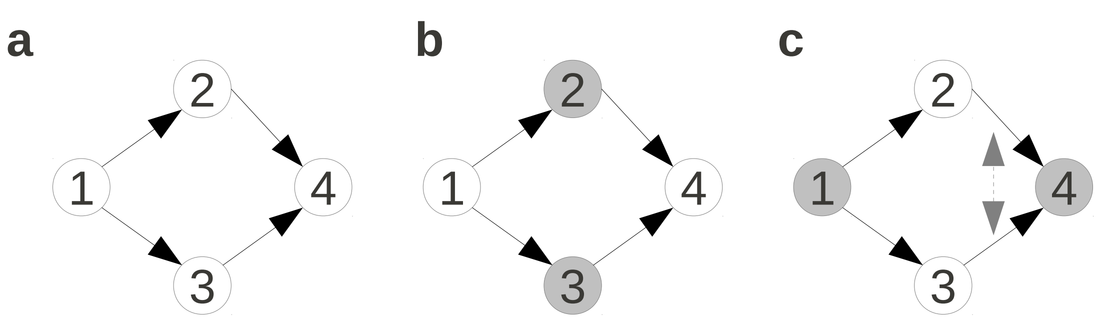


Figure a shows the model without conditioning. b shows $X_1 \perp X_4 | \{X_{2,} X_3\}$, and no Baye's balls can pass from 1 to 4. However, c shows conditioning on $\{X_{1,} X_4\}$ and allows a Baye's ball to pass through the “explaining away” subgraph on $\{X_{2,} X_3 X_4\}$, illustrated by the gray bidirectional arrow.

(b) The statement to prove/disprove is:

$$X_1 \perp X_2 | X_4 \text{ and } X_1 \perp X_3 | X_4 \text{ implies } X_1 \perp \{X_{2,} X_3\} | X_4$$

This statement can be shown as false, the intuition being that pairwise conditional independence does not imply joint conditional independence. A counterexample can be provided, by letting X_i be a binary random variable, and define each i=1,2,3,4 as:

$$X_1 = 1$$

$$X_2 = X_3 = Binary\,Coin\,Toss\,with\;p = \frac{1}{2}$$

$$X_4 = X_1 (X_2 + X_3) \text{ with modulo 2 addition}$$

To prove that the joint conditional independence implication is violated, we need to show that:

$$P(X_{1,} X_{2,} X_3 | X_4) \neq P(X_1 | X_4) P(X_{2,} X_3 | X_4)$$

The joint conditional probabilities are given as follows:

$(X_{1,} X_{2,} X_3)$	$P(X_{1,} X_{2,} X_3 \vert X_4 = 0)$
0, 0, 0	1/6
0, 0, 1	1/6
0, 1, 0	1/6
0, 1, 1	1/6
1, 0, 0	1/6
1, 0, 1	0
1, 1, 0	0
1, 1, 1	1/6

$(X_{1,} X_{2,} X_3)$	$P(X_{1,} X_{2,} X_3 \vert X_4 = 1)$
0, 0, 0	0
0, 0, 1	0
0, 1, 0	0
0, 1, 1	0
1, 0, 0	0
1, 0, 1	1/2
1, 1, 0	1/2
1, 1, 1	0

$(X_{2,} X_3)$	$P(X_{2,} X_3 \vert X_4 = 0)$
0, 0	1/3
0, 1	1/6
1, 0	1/6
1, 1	1/3

$(X_{2,} X_3)$	$P(X_{2,} X_3 \vert X_4 = 1)$
0, 0	0
0, 1	1/2
1, 0	1/2
1, 1	0

Evaluating at $(X_{1,} X_{2,} X_3) = (0,0,0)$ gives:

$$P(X_{1,} X_{2,} X_3 = (0,0,0) | X_4 = 0) = 1/6$$

$$P(X_1 = 0 | X_4 = 0) = 2/3 \quad , \quad P(X_{1,} X_2 = (0,0) | X_4 = 0) = 1/3$$

$$P(X_1 = 0 | X_4 = 0) P(X_{1,} X_2 = (0,0) | X_4 = 0) = 2/9$$

so that $P(X_{1,} X_{2,} X_3 = (0,0,0) | X_4 = 0) \neq P(X_1 = 0 | X_4 = 0) P(X_{1,} X_2 = (0,0) | X_4 = 0)$, the joint conditional independence assumption is violated, and the original implication is false.

(c) The statement to prove/disprove is:

$$X_1 \perp \{X_{2,} X_3\} | X_4 \quad \text{implies} \quad X_1 \perp X_2 | X_4$$

Ran out of time!

(d) Show these statements are equivalent: $(i)\, X \perp Y | Z$, $(ii)\, p(x, y, z) = h(x, z) k(y, z)$, $(iii)\, p(x, y, z) = p(x, z) p(y, z) / p(z)$.

$(i) \Rightarrow (iii)$:

$p(x, y, z) = p(x, y | z) p(z)$

$p(x, y, z) = p(x | z) p(y | z) p(z)$ by (i)

$p(x, y, z) = p(x, z) p(y, z) / p(z)$

$(i) \Rightarrow (ii)$:

$p(x, y | z) = p(x | z) p(y | z)$

$p(x, y, z) = p(x | z) p(z) p(y | z)$

Let $h(x, z) = p(x | z) p(z), k(y, z) = p(y | z)$

$(iii) \Rightarrow (i)$:

$p(x, z) p(y, z) / p(z) = p(x | z) p(z) p(y | z) p(z) / p(z)$

$p(x, z) p(y, z) / p(z) = p(x | z) p(y | z) p(z)$

$(iii) \Rightarrow (ii)$:

Let $h(x, z) = p(x, z) / p(z), k(y, z) = p(y, z)$

$(ii) \Rightarrow (i)$:

$p(x, y, z) = p(x | y, z) p(y | z) p(z)$

If $p(x, y, z) = h(x, z) k(y, z)$ then

$p(x | y, z) = p(x | z)$

$p(x, y, z) / p(z) = p(x | z) p(y | z)$

$p(x, y | z) = p(x | z) p(y | z)$

$(ii) \Rightarrow (iii)$:

$p(x, y, z) = p(x | y, z) p(y | z) p(z)$

$p(x, y, z) = p(x | z) p(y | z) p(z)$

Problem 4

(a) No, the ordering {1, …, 10} is not topological, because edge (10, 3) exists but 3 comes before 10 in the sort. One topological ordering for this graph is {1, 2, 6, 10, 3, 4, 7, 8, 9, 5}.

(b) Standard factorization (let $i = X_i$):

$$p(1...10) = p(5|9)\,p(9|3,7,8)\,p(8)\,p(7)\,p(4|2,7)\,p(3|10)\,p(10|2)\,p(6|1,2)\,p(2)\,p(1)$$

(c) For which (i, j) does $X_i \perp X_j$ hold?

(1,2), (1, 7), (1, 8), (1, 10), (1, 9), (1, 3), (1, 5)

(2, 1), (2, 7), (2, 8)

(3, 1), (3, 7), (3, 8), (3, 5)

(4, 8)

(5, 1)

(6, 7), (6, 8)

(7, 1), (7, 2), (7, 3), (7, 6), (7, 10), (7, 8)

(8, 1), (8, 2), (8, 3), (8, 4), (8, 6), (8, 7), (8, 10)

(9, 1)

(10, 1), (10, 7), (10, 8)

(d) What is the largest set A such that $X_1 \perp X_A | \{X_{2,} X_9\}$ holds?

A = {10, 3, 7, 8, 5}

(e) What is the largest set B where $X_8 \perp X_B | \{X_{2,} X_9\}$?

B = {1, 6, 5}

(f) How to draw from $p(X_5)$ without computing marginal?

I'm not sure?!? Maybe it involves using the Elimination algorithm.

Problem 5

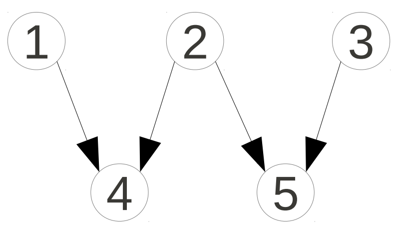
Given $X_i = Coin\,Toss\,Lands\,on\,Heads$, $p(X_i)=q$, $i \in \{1,2,3\}$, and $Z_4 = X_1 + X_2,\, Z_5 = X_2 + X_3$ with modulo 2 addition.

(a)

(X_2, X_3)	$p(X_2, X_3 \mid Z_5=0)$
0, 0	$q-q^2$
0, 1	0
1, 0	0
1, 1	$q-q^2$

(X_2, X_3)	$p(X_2, X_3 \mid Z_5=1)$
0, 0	0
0, 1	$q-q^2$
1, 0	$q-q^2$
1, 1	0

(b)



$X_1 \perp X_2,\, X_1 \perp X_3,\, X_1 \perp Z_4,\, X_1 \perp Z_5$

$X_2 \perp X_3,\, X_2 \perp Z_4,\, X_2 \perp Z_5$

$X_3 \perp Z_4,\, X_3 \perp Z_5$

$X_1 \perp X_2 \perp X_3$

(c)

Not sure how to draw an undirected graphical model for these relationships. I don't see how conditional *dependence* relationships for $X_1, X_2 \mid Z_4$ and $X_2, X_3 \mid Z_5$ can be represented in an undirectional way.

(d)

First we'll compute the marginal distribution for Z_4 and conditional distribution for $(X_1, Z_4) \mid X_2$:

$$p(Z_4=0) = p(X_1=1, X_2=1) + P(X_1=0, X_2=0) = 2q^2 - 2q + 1$$

$$p(Z_4=1) = p(X_1=1, X_2=0) + P(X_1=0, X_2=1) = 2q - 2q^2$$

(X_1, Z_4)	$p(X_1, Z_4 \mid X_2=0)$
(0, 0)	$(1-q)$
(0, 1)	0
(1, 0)	0
(1, 1)	q

(X_1, Z_4)	$p(X_1, Z_4 \mid X_2=1)$
(0, 0)	0
(0, 1)	$(1-q)$
(1, 0)	q
(1, 1)	0

The joint distribution $p(X_1, Z_4) = p(X_1, Z_4 \mid X_2=0)\, p(X_2=0) + p(X_1, Z_4 \mid X_2=1)\, p(X_2=1)$.

Independence implies $p(X_{1,}Z_4)=p(X_1)p(Z_4)$. Given an example of $(X_{1,}Z_4)$, such as $(X_{1,}Z_4)=(0,0)$, the computation of the joint distribution implies:

$$p(X_1=0,Z_4=0)=p(X_1=0,Z_4=0|X_2=0)p(X_2=0)+p(X_1=0,Z_4=0|X_2=1)p(X_2=1)$$

$$p(X_1=0,Z_4=0)=(1-q)(1-q)$$

Setting this equal to $p(X_1=0)p(Z_4=0)=(1-q)(2q^2-2q+1)$ gives:

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

$$1-q=2q^2-2q+1$$

$$2q^2-q=0$$

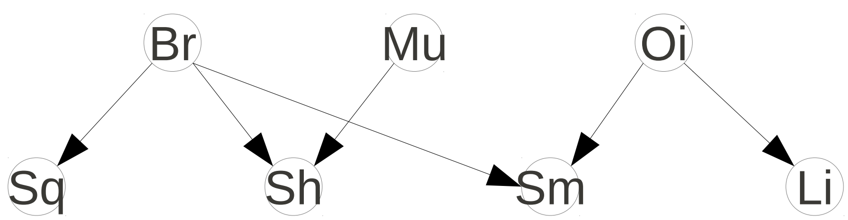
$$2q-1=0$$

$$q=1/2$$

This shows a value of q=1/2 or q=0 implies $X_1 \perp Z_4$.

Problem 6

(a)



(b)

Given Li, $P(Br|Li)=P(Li|Br)P(Br)/P(Li)$, since $P(Li|Br)=0$, $P(Br|Li)=0$. By a similar argument, $P(Mu|Li)=0$. However, $P(Oi|Li)=P(Li|Oi)P(Oi)/P(Li)\neq 0$. An engine light gives information only about an oil problem.

(c) Given Sm, Sh, what does Li tell us about Br or Mu?

We have $P(Br|Sm,Sh)=P(Sm,Sh|Br)P(Br)/P(Sm,Sh)$ and $P(Br|Sm,Sh,Li)=P(Sm,Sh,Li|Br)P(Br)/P(Sm,Sh,Li)$. It is not necessarily the case that $P(Sm,Sh,Li|Br)/P(Sm,Sh,Li)=P(Sm,Sh|Br)/P(Sm,Sh)$, so the engine light could reduce the probability of a brake problem or increase it. The same argument applies to a muffler problem.

(d) Given Br, a Baye's ball cannot pass between Sm and Mu, so $Sm \perp Mu|Br$, measuring Sm is irrelevant for determining a muffler problem.