Homework 2

Exercise 1. Consider a Bernoulli random variable X with success probability p. Suppose that we consider the following function: for every value X takes, let the transformed value be minus the log of the probability, basically the loglikelihood (with a minus sign, since the log of a probability is negative). Write the expected value of this transformed variable. ∎

$$g(x) = -\ln P(X = x)$$

$$E[g(X)] = \sum_x -P(X = x) \ln P(X = x)$$

$$E[g(X)] = -P(X = 0) \ln P(X = 0) - P(X = 1) \ln P(X = 1)$$

$$E[g(X)] = -(1 - p) \ln(1 - p) - p \ln p$$

Exercise 2. The topological sort is not unique. Write down a different topologically sorted ordering of the variables, and apply the same chain rule. Did it make any difference? ∎

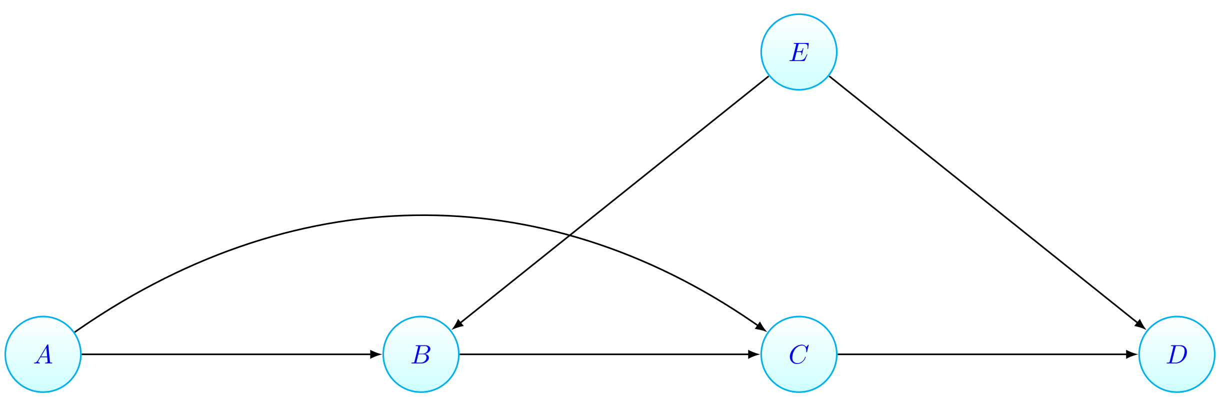
Answers may vary.

Topological sort: $X_0 X_1 X_2 Y_0 Y_1 Y_2$

$$P(X_0, X_1, X_2, Y_0, Y_1, Y_2) = P(Y_2|X_0, X_1, X_2, Y_0, Y_1) P(Y_1|X_0, X_1, X_2, Y_0) P(Y_0|X_0, X_1, X_2) \times P(X_2|X_0, X_1) P(X_1|X_0) P(X_0)$$

$$P(X_0, X_1, X_2, Y_0, Y_1, Y_2) = P(Y_2|X_2) P(Y_1|X_1) P(Y_0|X_0) P(X_2|X_1) P(X_1|X_0) P(X_0)$$

Exercise 3.

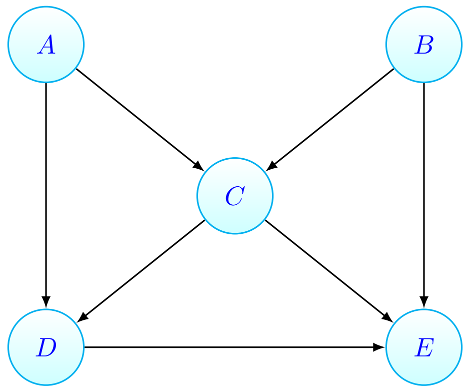


The graph above is called a *Verma graph*. Use the chain rule for Bayes nets to write the full joint distribution in terms of conditional probabilities. ∎

Topological sort: $AEBCD$

$$P(A,B,C,D,E) = P(D|C,E)P(C|A,B)P(B|A)P(E)P(A)$$

Exercise 4.



Use the chain rule for Bayes nets to write the full joint distribution in terms of conditional probabilities. ∎

Topological sort: $ABCDE$

$$P(A,B,C,D,E) = P(E|B,C,D)P(D|A,C)P(C|A,B)P(A)P(B)$$

Exercise 5. Show $\bar{y} = p/(1-(1-p)(1-b))$. If $0 < r < 1$, what is $\lim_{\tau\to\infty} r^\tau$? (Just the answer intuitively, not a formal proof.) Use your formula from the previous problem and show that as $\tau \to \infty$, $y_\tau \to \bar{y}$. This is our first example of something quite important: the convergence of a Markov chain to a stationary distribution. ∎

There are three parts to this problem:

1)

$$y_{\tau+1} = p + ky_\tau$$

$$\bar{y} = p + k\bar{y}$$

$$\bar{y} = p/(1-k)$$

where $k = (1-p)(1-b)$.

2)

$$\lim_{\tau->\infty} r^\tau = 0$$

3)

From before, we have:

$$y_\tau = \frac{p(1-k^\tau)}{1-k}$$

$$\lim_{\tau->\infty} y_\tau = \lim_{\tau->\infty} \frac{p(1-k^\tau)}{1-k} = \frac{p}{1-k} \lim_{\tau->\infty}(1-k^\tau) = \frac{p}{1-k},$$

where $k = (1-p)(1-b)$.

Exercise 6. Consider the value of $\bar{y}$ for these extreme special cases: $p = 0$, $p = 1$, $b = 0$, $b = 1$. Does the formula give interpretable results? What if $p = b = 1$? ∎

$\bar{y} = p/(1-(1-p)(1-b))$

Case	$\bar{y}$
$p = 0$	0
$p = 1$	1
$b = 0$	1
$b = 1$	p
$p = b = 1$	1

Exercise 7. Suppose that in a certain region, the prevalence Q of infection is either 10%, 20%, or 40% in a district, and that the probabilities are 40%, 50%, and 10%, respectively. (In other words, forty percent of the districts have a prevelance of 10%, fifty percent have a prevalence of 20%, and ten percent have a prevalence of 40%.) In each district, suppose you sample N children and estimate the number X with the infection. Assume simple random sampling; given the prevalence, the number of infected children in the sample is going to be assumed binomial (the districts are much larger than the sample size).

Prevalence π	$P(Q = \pi)$	$E[X\|Q = \pi]$
0.1	0.4	
0.2	0.5	
0.4	0.1	

The quantity $E[X|Q]$ is a random variable taking what values with what probabilities? Find the expected value of the random variable $E[X|Q]$. ∎

Prevalence π	$P(Q = \pi)$	$E[X\|Q = \pi]$
0.1	0.4	$0.1N$
0.2	0.5	$0.2N$
0.4	0.1	$0.4N$

$P(E[X|Q] = 0.1N) = 0.4$

$P(E[X|Q] = 0.2N) = 0.5$

$P(E[X|Q] = 0.4N) = 0.1$

$E[E[X|Q]] = 0.1N * 0.4 + 0.2N * 0.5 + 0.4N * 0.1 = 0.18N$

Exercise 8. Show the conditional variance of Y given $X = 1$ is 900 in our example. ∎

$var(Y|X = 1) = E[Y^2|X = 1] - (E[Y|X = 1])^2$

$E[Y|X = 1] = 10 \times 0.72/0.8 + 110 \times 0.08/0.8 = 20$

$E[Y^2|X = 1] = 10^2 \times 0.72/0.8 + 110^2 \times 0.08/0.8 = 1300$

$var(Y|X = 1) = 1300 - 20^2 = 900$

Exercise 9. How long did this assignment take you? ∎