Computational Statistics
Homework 1

1. (10 points) Suppose that a regular light bulb, a long-life light bulb, and an extra-long-life light bulb are being tested. The lifetime X_1 of the regular bulb has the exponential distribution with mean θ, the lifetime X_2 of the long-life bulb has the exponential distribution with mean 2θ, and the lifetime X_3 of the extra-long-life bulb has the exponential distribution with mean 3θ. *Reminder: The exponential distribution has density* $f(x) = \lambda e^{-\lambda x}, x > 0$ *and* $E(X) = \frac{1}{\lambda}$. Determine the M.L.E. of θ based on the observations x_1, x_2, x_3.

2. (10 points) The Pareto distribution with parameter θ has range $[1, \infty)$ and pdf: Η κατανομή Pareto με παράμετρο θ έχει εύρος $[1, \infty)$ και σππ:
$$f(x) = \frac{\theta}{x^{\theta+1}}$$
Suppose the data $5, 2, 3$ was drawn independently from such a distribution. Assume that you are Bayesian, and your prior for θ is a Gamma distribution with parameters α, β. Find the posterior distribution for θ. *Reminder: The pdf of the Gamma distribution with parameters* α, β: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$

3. (10 points) Consider the joint probability
$$p(A, B, C, D) = p(A \mid B)p(B \mid C)p(D \mid C)P(C).$$
   i Draw the corresponding Bayesian network.
   ii Consider B and D; which latent variables make them conditionally independent? Prove it explicitly.

4. (10 points) Let $\phi(A, B, C)$ be a factor in a probability distribution that factorizes over a Markov network. Which of the following must be true? You may select 1 or more options.
   1. There is a path connecting A, B, and C in the network.
   2. $\phi(a, b, c) \geq 0$, where a is a value of A, b is a value of B, and c is a value of C.
   3. A, B, and C form a clique in the network.
   4. $\phi(a, b, c) \leq 1$, where a is a value of A, b is a value of B, and c is a value of C.
   5. A, B, and C do not form a clique in the network.

5. (10 points) Consider the probabilistic model linking economy (E) to inflation (I) and price of oil (O) through observations of stock prices (S) and prices of oil futures (F). Suppose that all the quantities can take on two distinct values of high (h) and low (l), except for the stock price which can also be normal (n).

$p(E = l) = 0.2$	
$p(S = l \mid O = l) = 0.9$	$p(S = n \mid O = l) = 0.1$
$p(S = l \mid O = h) = 0.1$	$p(S = n \mid O = h) = 0.4$
$p(O = l \mid E = l) = 0.9$	$p(O = l \mid E = h) = 0.05$
$p(F = l \mid I = l, E = l) = 0.9$	$p(F = l \mid I = l, E = h) = 0.1$
$p(F = l \mid I = h, E = l) = 0.1$	$p(F = l \mid I = h, E = h) = 0.01$
$p(I = l \mid O = l, E = l) = 0.9$	$p(I = l \mid O = l, E = h) = 0.1$
$p(I = l \mid O = h, E = l) = 0.1$	$p(I = l \mid O = h, E = h) = 0.01$

   i Using the conditional probabilities above, draw the Bayesian network corresponding to this model.
   ii Then, compute the probability that inflation (I) is high, given that the stock price (S) is normal and the futures price (F) is high.

6. (20 points) Implement Algorithm 10.A.1 in page 259 of the book "Probabilistic Graphical Models" by D. Koller and N. Friedman.

7. (20 points) For the graph in Figure 1, run (by hand) all the steps of the variable elimination algorithm for some ordering you pick.

8. (10 points) Let G be a Bayesian network with no v-structures. Let H be a Markov network with the same skeleton as G. Show that H is an I-Map $I_D(G)$.

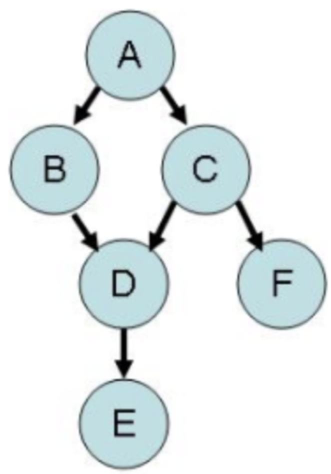

Figure 1: Markov Network for Ex. 7