MIDTERM EXAM

ISyE6420
Fall 2020

Released October 16, 12:00pm – due October 25, 11:55pm. This exam is not proctored and not time limited except the due date. Late submissions will not be accepted.

Use of all available electronic and printed resources is allowed except direct communication that violates Georgia Tech Academic Integrity Rules.

Name ____________________

Problem	1	2	3	Total
Score	/33	/33	/34	/100

1. Bayes Network. Incidences of diseases A and B (D_A, D_B) depend on the exposure (E). Disease A is additionally influenced by risk factors (R). Both diseases lead to symptoms (S). Results of the test for disease A (T_A) are affected also by disease B. Positive test will be denoted as $T_A = 1$, negative as $T_A = 0$. The Bayes Network is shown in Figure 1. Needed conditional probabilities are shown in Table 1.

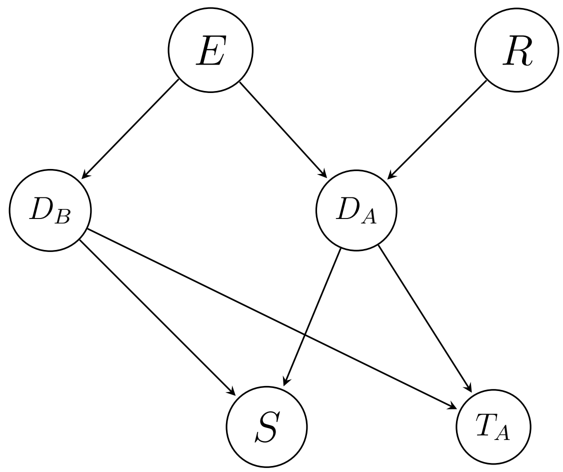

Figure 1: The DAG of the Bayesian networks

Table 1: The known (or elicited) conditional probabilities

E	0	1
	0.8	0.2

R	0	1
	0.7	0.3

D_A	0	1
E^cR^c	0.9	0.1
E^cR	0.4	0.6
ER^c	0.5	0.5
ER	0.3	0.7

D_B	0	1
E^c	0.8	0.2
E	0.3	0.7

S	0	1
$D_A^cD_B^c$	0.95	0.05
$D_A^cD_B$	0.6	0.4
$D_AD_B^c$	0.4	0.6
D_AD_B	0.1	0.9

T_A	0	1
$D_A^cD_B^c$	0.92	0.08
$D_A^cD_B$	0.8	0.2
$D_AD_B^c$	0.15	0.85
D_AD_B	0.03	0.97

(a) What is the probability of disease A ($D_A = 1$), if disease B is not present ($D_B = 0$), but symptoms are present ($S = 1$).

(b) What is the probability of exposure ($E = 1$), if symptoms are present ($S = 1$) and test is positive ($T_A = 1$).

Hint: You can solve this problem by any of the 3 ways: (i) use of WinBUGS or Open-

BUGS, (ii) direct simulation using Octave/MATLAB, R, or Python, and (iii) exact calculation.

2. Times to Failure. Three devices are monitored until failure. The observed lifetimes are 0.9, 1.8, and 0.3 years. If the lifetimes ate modeled as exponential distribution with rate λ,

$$T_i \sim \mathcal{E}xp(\lambda), \quad f(t|\lambda) = \lambda e^{-\lambda t}, t > 0, \lambda > 0.$$

Assume exponential prior on λ,

$$\lambda \sim \mathcal{E}xp(2), \quad \pi(\lambda) = 2e^{-2\lambda}, \lambda > 0.$$

(a) Find the posterior distribution of λ.

(b) Find the Bayes estimator for λ.

(c) Find the MAP estimator for λ.

(d) Numerically find 95% equitailed confidence interval for λ.

(e) Find the posterior probability of hypothesis $H_0 : \lambda \leq 1/2$.

3. Gibbs and High/Low Protein Diet in Rats. Armitage and Berry (1994, p. 111) report data on the weight gain of 19 female rats between 28 and 84 days after birth. The rats were placed on diets with high (12 animals) and low (7 animals) protein content.

High protein	Low protein
134	70
146	118
104	101
119	85
124	107
161	132
107	94
83	
113	
129	
97	
123	

We want to test the hypothesis on dietary effect. Did a low protein diet result in significantly lower weight gain?

The classical t test against one sided alternative will be significant. We will do the test Bayesian way using Gibbs sampler.

Assume that high-protein diet measurements $y_{1i}, i = 1, \ldots, 12$ are coming from normal distribution $\mathcal{N}(\theta_1, 1/\tau_1)$, where τ_1 is precision parameter,

$$f(y_{1i}|\theta_1, \tau_1) \propto \tau_1^{1/2} \exp\left\{-\frac{\tau_1}{2}(y_{1i} - \theta_1)^2\right\}, i = 1, \ldots, 12.$$

Low-protein diet measurements $y_{2i}, i = 1, \ldots, 7$ are coming from normal distribution $\mathcal{N}(\theta_2, 1/\tau_2)$,

$$f(y_{2i}|\theta_2, \tau_2) \propto \tau_2^{1/2} \exp\left\{-\frac{\tau_2}{2}(y_{2i} - \theta_2)^2\right\}, i = 1, \ldots, 7.$$

Assume that θ_1 and θ_2 have normal priors $\mathcal{N}(\theta_{10}, 1/\tau_{10})$ and $\mathcal{N}(\theta_{20}, 1/\tau_{20})$, respectively. Take prior means as $\theta_{10} = \theta_{20} = 110$ (apriori no preference) and precisions as $\tau_{10} = \tau_{20} = 1/100$.

Assume that τ_1 and τ_2 have the gamma $\mathcal{G}a(a_1, b_2)$ and $\mathcal{G}a(a_2, b_2)$ priors with shapes $a_1 = a_2 = 0.01$ and rates $b_1 = b_2 = 4$.

(a) Construct Gibbs sampler that will sample $\theta_1, \tau_1, \theta_2$, and τ_2 from their posteriors.

(b) Find sample differences $\theta_1 - \theta_2$. Proportion of positive differences approximates posterior probability of hypothesis $H_0 : \theta_1 > \theta_2$. What is this proportion?

(c) Using sample quantiles find the 95% equitailed credible set for $\theta_1 - \theta_2$. Does this set contain 0?