CS6957: Probabilistic Modeling **Prateep Mukherjee(u0876583)**

Homework 2

1

The total number of probabilities needed to store the network is product of the number of values for all the variables. Therefore, the result is $[8(income) * 2(exercise) * 2(smoke) * 4(bmi) * 4(bp) * 2(cholesterol) * 2(angina) * 2(stroke) * 2(attack) * 4(diabetes)] = 32768$.

2 (a)

p(Diabetes | bad habits)

probs	smoke	exercise	diabetes
0.136660053	1	2	1
0.008914666	1	2	2
0.837385140	1	2	3
0.017040141	1	2	4

p(Diabetes | good habits)

probs	smoke	exercise	diabetes
0.131893738	2	1	1
0.008881042	2	1	2
0.842538131	2	1	3
0.016687089	2	1	4

p(Stroke | bad habits)

probs	smoke	exercise	stroke
0.04959854	1	2	1
0.95040146	1	2	2

p(Stroke | good habits)

probs	smoke	exercise	stroke
0.03605402	2	1	1
0.96394598	2	1	2

p(Attack | bad habits)

probs	smoke	exercise	attack
0.0742565	1	2	1
0.9257435	1	2	2

p(Attack | good habits)

probs	smoke	exercise	attack
0.05287024	2	1	1
0.94712976	2	1	2

p(Angina | bad habits)

probs	smoke	exercise	angina
0.08008408	1	2	1
0.91991592	1	2	2

p(Angina | good habits)

probs	smoke	exercise	angina
0.05488682	2	1	1
0.94511318	2	1	2

(b)

p(Diabetes | poor health)

probs	diabetes	bmi	cholesterol	bp
0.115422719	1	3	1	1
0.007661825	2	3	1	1
0.860872761	3	3	1	1
0.016042695	4	3	1	1

p(Diabetes | good health)

probs	diabetes	bmi	cholesterol	bp
0.057709954	1	2	2	3
0.009543386	2	2	2	3
0.922193878	3	2	2	3
0.010552782	4	2	2	3

p(Stroke | poor health)

probs	stroke	bmi	bp	cholesterol
0.08268577	1	3	1	1
0.91731423	2	3	1	1

p(Stroke | good health)

probs	stroke	bmi	bp	cholesterol
0.01446014	1	2	3	2
0.98553986	2	2	3	2

p(Attack | poor health)

probs	attack	bmi	bp	cholesterol
0.1407844	1	3	1	1
0.8592156	2	3	1	1

p(Attack | good health)

probs	attack	bmi	bp	cholesterol
0.01616133	1	2	3	2
0.98383867	2	2	3	2

p(Angina | poor health)

probs	angina	bmi	bp	cholesterol
0.1616076	1	3	1	1
0.8383924	2	3	1	1

p(Angina | good health)

probs	angina	bmi	bp	cholesterol
0.01332601	1	2	3	2
0.98667399	2	2	3	2

3

From Fig. 1, the inference is that as a person's annual income level goes higher, the probability of all the health outcomes reduces.

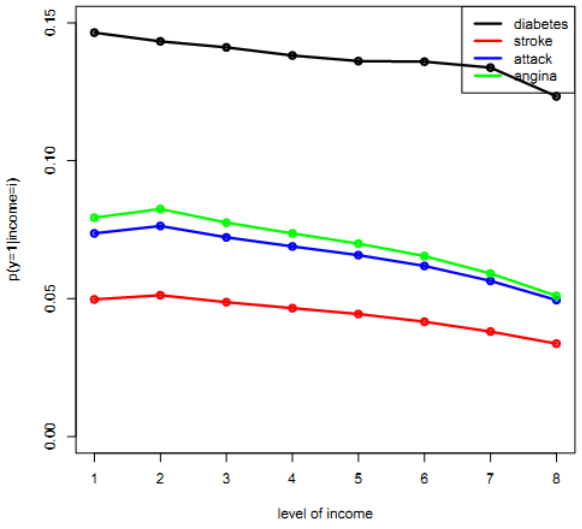

Figure 1: Probability of health outcome(diabetes, stroke, heart attack, angina) for each level of annual income. Annual income is plotted in 8 different levels. Posterior probability $P(y = 1|income = i)$ is normalized to [0,0.15].

4

As there are no links between the habits(smoking and exercise) and the outcomes, the assumption is that there are no direct effects of smoking and exercise on the diseases. Next, we add links between the habits and the outcomes. The following tables show the inference measures obtained.

p(Diabetes | bad habits)

probs	smoke	exercise	diabetes
0.210944859	1	2	1
0.006915095	1	2	2
0.760692694	1	2	3
0.021447352	1	2	4

p(Diabetes | good habits)

probs	smoke	exercise	diabetes
0.098552162	2	1	1
0.009884084	2	1	2
0.877575578	2	1	3
0.013988176	2	1	4

p(Diabetes | poor health)

probs	diabetes	bmi	cholesterol	bp
0.123480634	1	3	1	1
0.007460298	2	3	1	1
0.852415963	3	3	1	1
0.016643105	4	3	1	1

p(Diabetes | good health)

probs	diabetes	bmi	cholesterol	bp
0.054172949	1	2	2	3
0.009731215	2	2	2	3
0.925952333	3	2	2	3
0.010143502	4	2	2	3

p(Stroke | bad habits)

probs	smoke	exercise	stroke
0.07803498	1	2	1
0.92196502	1	2	2

p(Stroke | good habits)

probs	smoke	exercise	stroke
0.02431088	2	1	1
0.97568912	2	1	2

p(Stroke | poor health)

probs	stroke	bmi	bp	cholesterol
0.08425692	1	3	1	1
0.91574308	2	3	1	1

p(Stroke | good health)

probs	stroke	bmi	bp	cholesterol
0.01399739	1	2	3	2
0.98600261	2	2	3	2

p(Attack | bad habits)

probs	smoke	exercise	attack
0.1211659	1	2	1
0.8788341	1	2	2

p(Attack | good habits)

probs	smoke	exercise	attack
0.03101531	2	1	1
0.96898469	2	1	2

p(Attack | poor health)

probs	attack	bmi	bp	cholesterol
0.1421993	1	3	1	1
0.8578007	2	3	1	1

p(Attack | good health)

probs	attack	bmi	bp	cholesterol
0.01546893	1	2	3	2
0.98453107	2	2	3	2

p(Angina | bad habits)

probs	smoke	exercise	angina
0.1190069	1	2	1
0.8809931	1	2	2

p(Angina | good habits)

probs	smoke	exercise	angina
0.03680005	2	1	1
0.96319995	2	1	2

p(Angina | poor health)

probs	angina	bmi	bp	cholesterol
0.1629716	1	3	1	1
0.8370284	2	3	1	1

p(Angina | good health)

probs	angina	bmi	bp	cholesterol
0.01294426	1	2	3	2
0.98705574	2	2	3	2

The above tables show that probability for people having good habits and health for not having any health outcomes increase. Similarly, probability for people having health outcomes given they have poor habits and health increases. This shows that our assumption in question 5 is not valid.

5

The bayesian net in question assumes no direct link between the health outcomes. After we introduce such links, the probability that a person can get stroke given he already has diabetes increases. Similarly, probability that he can get stroke given that he does not have diabetes decreases.

P_{old}(stroke = 1 | diabetes=1)	P_{new}(stroke=1 | diabetes=1)
0.04416376	0.07619782

P_{old}(stroke = 1 | diabetes=3)	P_{new}(stroke=1 | diabetes=3)
0.04047831	0.03501533

The above tables show that the above assumption is not valid.