Concepts in probability
BMI 826-023 Fall 2020

Prof. Sushmita Roy

Problem 1.

This question should be helpful to walk through the steps of testing conditional independence using the example we saw in class. Recall the **Table 1** specifying the probability distribution for three random variables, *Fever*, *Flu*, *Headache*. Using this table, we would like to ask if *Fever* and *Headache* are independent or

Flu	*Fever*	*Headache*	$P(Flu, Fever, Headache)$
TRUE	TRUE	TRUE	0.04
TRUE	TRUE	FALSE	0.04
TRUE	FALSE	TRUE	0.01
TRUE	FALSE	FALSE	0.01
FALSE	TRUE	TRUE	0.009
FALSE	TRUE	FALSE	0.081
FALSE	FALSE	TRUE	0.081
FALSE	FALSE	FALSE7	0.729

Table 1: Joint distribution of three random variables

conditionally independent given *Flu*.

(a) Using **Table 1** compute the joint distribution, $P(Fever, Headache)$.

(b) Using **Table 1** compute the marginal distributions, $P(Fever)$ and $P(Headache)$.

(c) Using the distributions computed in (a) and (b), show that *Fever* and *Headache* are not independent, that is, $P(Fever, Headache) \neq P(Fever) * P(Headache)$

(d) Using **Table 1** compute the conditional distributions, $P(Fever, Headache|Flu)$, $P(Fever|Flu)$ and $P(Headache|Flu)$.

(e) Using the distributions computed in (d), show that *Fever* and *Flu* are conditionally independent given *Flu*. In particular show that $P(Fever, Headache|Flu) = P(Fever|Flu) * P(Headache|Flu)$

Problem 2.

Assume we are interested in the distribution of food items, namely, drinks and snacks ordered at a local grocery store. We are also interested in studying if there is any dependency between the food items purchased and the time of day. We will represent these food items and time of day using the random variables

D, S and T. D denotes a drink and takes on values $\{$*coffee*, *water*$\}$. S denotes a snack and takes on values $\{cereal, chips\}$. T denotes the time of day, and takes on values $\{$*morning*, *noon*$\}$. We are given the following set of observations for these events, each row in the table below representing an observation for D, S and T. Using these observations, please answer the questions (a)-(f) below.

Drink	*Snack*	*TimeOfDay*
coffee	chips	morning
coffee	cereal	morning
coffee	cereal	morning
coffee	cereal	morning
water	cereal	morning
water	chips	noon
coffee	chips	noon
water	cereal	noon
coffee	cereal	noon

(a) Estimate the probability distribution, $P(S)$

(b) Estimate the joint probability distribution, $P(S, D)$

(c) Estimate the marginal probability distribution of $P(S)$ from the joint, $P(S, D)$.

(d) Is *S* independent of *D*?

(e) Is *S* independent of *D* given T = *morning*?

(f) Is *S* independent of *D* given T = *noon*?