Foundations
Probability: Univariate Models

Wafaa Mohammed

Exercise 2.1. [Conditional independence]
a. Let $H \in 1, ..., K$ be a discrete random variable, and let e_1 and e_2 be the observed values of two other random variables E_1 and E_2. Suppose we wish to calculate the vector

$$\overrightarrow{P}(H|e_1, e_2) = (P(H = 1|e_1, e_2), ..., P(H = K|e_1, e_2))$$

Which of the following sets of numbers are sufficient for the calculation?

i. $P(e_1, e_2), P(H), P(e_1|H), P(e_2|H)$

ii. $P(e_1, e_2), P(H), P(e_1, e_2|H)$

iii. $P(e_1|H), P(e_2|H), P(H)$

b. Now suppose we now assume $E_1 \perp E_2|H$ (i.e., E_1 and E_2 are conditionally independent given H). Which of the above 3 sets are sufficient now?

Show your calculations as well as giving the final result. Hint: use Bayes rule.

Solution. a. $\overrightarrow{P}(H|e_1, e_2) = \frac{P(H)P(e_1,e_2|H)}{P(e_1,e_2)}$, set ii will be sufficient.

b. in case of conditional independence, $P(e_1, e_2|H) = P(e_1|h)P(e_2|h)$, so set i will be sufficient in this case.

□