Concepts in probability
BMI 826-023 Fall 2018

Sushmita Roy & Alireza Fotuhi Siahpirani

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 values {*morning*, *noon*}. We are the given the following set of observations for these events, each row in the table below representing a combination 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*?

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

$|S = chips| = 3$
$|S = cereal| = 6$
$P(S = chips) = \frac{1}{3}$
$P(S = cereal) = \frac{2}{3}$

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

$|S = chips\&D = coffee| = 2$
$|S = cereal\&D = coffee| = 4$
$|S = chips\&D = water| = 1$
$|S = cereal\&D = water| = 2$
$P(S = chips\&D = coffee) = \frac{2}{9}$
$P(S = cereal\&D = coffee) = \frac{4}{9}$
$P(S = chips\&D = water) = \frac{1}{9}$
$P(S = cereal\&D = water) = \frac{2}{9}$

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

$\sum_d P(S = chips\&D = d) = \frac{2}{9} + \frac{1}{9} = \frac{1}{3}$
$\sum_d P(S = cereal\&D = d) = \frac{4}{9} + \frac{2}{9} = \frac{2}{3}$

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

Yes.
$|D = coffee| = 6$
$|D = water| = 3$
$P(D = coffee) = \frac{2}{3}$
$P(D = water) = \frac{1}{3}$
$P(S = chips\&D = coffee) = \frac{2}{9} = \frac{1}{3} \times \frac{2}{3} = P(S = chips)P(D = coffee)$
$P(S = cereal\&D = coffee) = \frac{4}{9} = \frac{2}{3} \times \frac{2}{3} = P(S = cereal)P(D = coffee)$
$P(S = chips\&D = water) = \frac{1}{9} = \frac{1}{3} \times \frac{1}{3} = P(S = chips)P(D = water)$
$P(S = cereal\&D = water) = \frac{2}{9} = \frac{2}{3} \times \frac{1}{3} = P(S = cereal)P(D = water)$

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

No.
$|S = chips\&T = morning| = 1$
$|S = cereal\&T = morning| = 4$
$|D = coffee\&T = morning| = 4$
$|D = water\&T = morning| = 1$
$|S = chips\&D = coffee\&T = morning| = 1$

$|S = cereal \& D = coffee \& T = morning| = 3$
$|S = chips \& D = water \& T = morning| = 0$
$|S = cereal \& D = water \& T = morning| = 1$
$P(S = chips \& D = water|T = morning) = \frac{0}{5} \neq \frac{1}{5} \times \frac{1}{5} = P(S = chips|T = morning)P(D = water|T = morning)$
$P(S = chips \& D = coffee|T = morning) = \frac{1}{5} \neq \frac{1}{5} \times \frac{4}{5} = P(S = chips|T = morning)P(D = coffee|T = morning)$
$P(S = cereal \& D = water|T = morning) = \frac{1}{5} \neq \frac{1}{5} \times \frac{4}{5} = P(S = cereal|T = morning)P(D = water|T = morning)$
$P(S = cereal \& D = coffee|T = morning) = \frac{3}{5} \neq \frac{4}{5} \times \frac{4}{5} = P(S = cereal|T = morning)P(D = coffee|T = morning)$

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

Yes.
$|S = chips \& T = noon| = 2$
$|S = cereal \& T = noon| = 2$
$|D = coffee \& T = noon| = 2$
$|D = water \& T = noon| = 2$
$|S = chips \& D = coffee \& T = noon| = 1$
$|S = cereal \& D = coffee \& T = noon| = 1$
$|S = chips \& D = water \& T = noon| = 1$
$|S = cereal \& D = water \& T = noon| = 1$
$P(S = chips \& D = water|T = noon) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(S = chips|T = noon)P(D = water|T = noon)$
$P(S = chips \& D = coffee|T = noon) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(S = chips|T = noon)P(D = coffee|T = noon)$
$P(S = cereal \& D = water|T = noon) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(S = cereal|T = noon)P(D = water|T = noon)$
$P(S = cereal \& D = coffee|T = noon) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(S = cereal|T = noon)P(D = coffee|T = noon)$