

# Probability [3]

Qi Zhao

July 12, 2018

## 1. Joint Probability [2]

Consider two random variables,  $x$  and  $y$ . A joint probability distribution may relate variables that are all discrete or all continuous, or it may relate discrete variables to continuous ones (see Figure 1). This information is encompassed in the joint probability distribution of  $x$  and  $y$ , which is written as  $\Pr(x,y)$ . The comma in  $\Pr(x,y)$  can be read as the English word and so  $\Pr(x,y)$  is the probability of  $x$  and  $y$ . And it writes  $\Pr(x,y,z)$  to represent the joint probability distribution of scalar variables  $x,y$ , and  $z$  or writes  $\Pr(x)$  to represent the joint probability of all of the elements of the multidimensional variable  $x = [x_1, x_2, \dots, x_K]^T$ . Finally, we will write  $\Pr(x,y)$  to represent the joint distribution of all of the elements from multidimensional variables [1]  $x$  and  $y$ . Regardless, the total probability of all outcomes (summing over discrete variables and integrating over continuous ones) is always one.

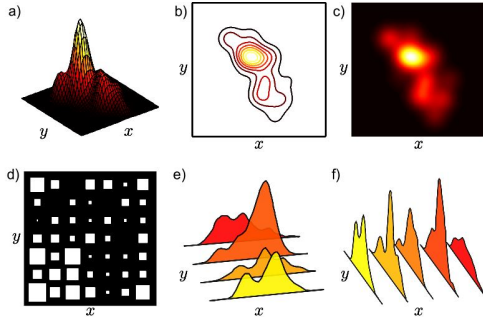


Figure 1. Joint probability distributions between variables  $x$  and  $y$ . a-c) The same joint pdf of two continuous variables represented as a surface, contour plot, and image, respectively. d) Joint distribution of two discrete variables represented as a 2D Hinton diagram. e) Joint distribution of a continuous variable  $x$  and discrete variable  $y$ . f) Joint distribution of a discrete variable  $x$  and continuous variable  $y$ .

## 2. Marginalization [4]

It can be recovered the probability distribution of any single variable from a joint distribution by summing (discrete

case) or integrating (continuous case) over all the other variables (Figure 2). And the distributions  $\Pr(x)$  and  $\Pr(y)$  using the relations in Equation 1 are recovered:

$$\begin{aligned} \Pr(x) &= \int \Pr(x,y) dy \\ \Pr(y) &= \int \Pr(x,y) dx \end{aligned} \quad (1)$$

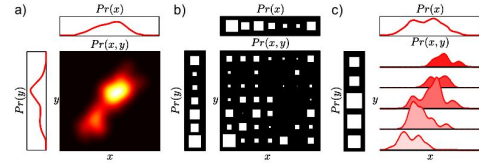


Figure 2. Joint and marginal probability distributions. The marginal probability  $\Pr(x)$  is found by summing over all values of  $y$  (discrete case) or integrating over  $y$  (continuous case) in the joint distribution  $\Pr(x,y)$ . Similarly, the marginal probability  $\Pr(y)$  is found by summing or integrating over  $x$ . Note that the plots for the marginal distributions have different scales from those for the joint distribution (on the same scale, the marginals would look larger as they sum all of the mass from one direction). a) Both  $x$  and  $y$  are continuous. b) Both  $x$  and  $y$  are discrete. c) The random variable  $x$  is continuous and the variable  $y$  is discrete.

In general, it can be recovered the joint probability of any subset of variables, by marginalizing over all of the others.

## References

- [1] M. Anderson and T. Adali. Second and higher-order correlation analysis of multiple multidimensional variables by joint diagonalization. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 197–204, 2010. 1
- [2] R. F. Carsel and R. S. Parrish. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5):755–769, 1988. 1
- [3] J. L. Snell. *Convergence of probability measures*. Wiley, 1968. 1

- [4] R. S. Strauss and H. A. Pollack. Social marginalization of overweight children. *Archives of Pediatrics & Adolescent Medicine*, 157(8):746–752, 2003. [1](#)