imply (absolute) independence, that is,

$$p(x, y \mid z) \;=\; p(x \mid z)\; p(y \mid z) \quad \not\Rightarrow \quad p(x, y) \;=\; p(x)\; p(y) \tag{2.20}$$

The converse is also in general untrue: absolute independence does not imply conditional independence:

$$p(x, y) \;=\; p(x)\; p(y) \quad \not\Rightarrow \quad p(x, y \mid z) \;=\; p(x \mid z)\; p(y \mid z) \tag{2.21}$$

In special cases, however, conditional and absolute independence may coincide.

EXPECTATION OF A RV

A number of probabilistic algorithms require us to compute features, or statistics, of probability distributions. The *expectation* of a random variable $X$ is given by

$$E[X] \;=\; \sum_x x\; p(x) \qquad \text{(discrete)} \tag{2.22}$$

$$E[X] \;=\; \int x\; p(x)\; dx \quad \text{(continuous)} \tag{2.23}$$

Not all random variables possess finite expectations; however, those that do not are of no relevance to the material presented in this book.

The expectation is a linear function of a random variable. In particular, we have

$$E[aX + b] \;=\; aE[X] + b \tag{2.24}$$

for arbitrary numerical values $a$ and $b$. The covariance of $X$ is obtained as follows

$$\mathrm{Cov}[X] \;=\; E[(X - E[X])^2] \;=\; E[X^2] - E[X]^2 \tag{2.25}$$

The covariance measures the squared expected deviation from the mean. As stated above, the mean of a multivariate normal distribution $\mathcal{N}(x; \mu, \Sigma)$ is $\mu$, and its covariance is $\Sigma$.

ENTROPY

A final concept of importance in this book is *entropy*. The entropy of a probability distribution is given by the following expression:

$$H_p(x) \;=\; E[-\log_2 p(x)] \tag{2.26}$$

which resolves to

$$H_p(x) \;=\; -\sum_x p(x)\; \log_2 p(x) \qquad \text{(discrete)} \tag{2.27}$$

$$H_p(x) \;=\; -\int p(x)\; \log_2 p(x)\; dx \quad \text{(continuous)} \tag{2.28}$$