Exercise 2.1 Probabilities are sensitive to the form of the question that was used to generate the answer

(a)

C1	C2
B	B
B	G
G	B
G	G

If the neighbor has any boys, there is only a chance to the first of the three rows in the above table. In this situation, the probability that one child is a girl is $\frac{2}{3}$.

(b)

$$
\begin{aligned}
P(Run = C1) \cdot P(C2 = G|Run = C1) + P(Run = C2) \cdot P(C1 = G|Run = C2) &= \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{1}{2}
\end{aligned}
$$

Exercise 2.2 Legal reasoning

(a)

$$
\begin{aligned}
G &= \begin{cases} 1 & \text{(defendant is guilty)} \\ 0 & \text{(otherwise)} \end{cases} && (1) \\
B &= \begin{cases} 1 & \text{(defendant's blood type matches one at the scene)} \\ 0 & \text{(otherwise)} \end{cases} && (2) \\
P(G = 0|B = 1) &= \frac{P(B = 1|G = 0) \cdot P(G = 0)}{P(B = 1)} && (3) \\
&= \frac{P(B = 1|G = 0) \cdot P(G = 0)}{P(B = 1|G = 1) \cdot P(G = 1) + P(B = 1|G = 0) \cdot P(G = 0)} && (4) \\
&= \frac{\frac{1}{1000} \times \frac{799999}{800000}}{1 \times \frac{1}{800000} + \frac{1}{1000} \times \frac{799999}{800000}} && (5) \\
&\simeq 0.999 && (6)
\end{aligned}
$$

The probability of the defendant is innocense given that defendant's blood type matches one at the scene is 0.999.

(b)

The probabilities of guilty of 8000 people are all the same. This fact doesn't show that the defendant is innocense.

Exercise 2.3 Variance of a sum

$$
\begin{aligned}
V[X+Y] &= E[\{(X+Y) - E[X+Y]\}^2] && (7)\\
&= E[\{(X+Y) - (E[X] + E[Y])\}^2] && (8)\\
&= E[\{(X - E[X]) + (Y - E[Y])\}^2] && (9)\\
&= E[(X - E[X])^2 + (Y - E[Y])^2 + 2(X - E[X])(Y - E[Y])] && (10)\\
&= E[(X - E[X])^2] + E[(Y - E[Y])^2] + 2E[(X - E[X])(Y - E[Y])] && (11)\\
&= V[X] + V[Y] + 2\mathrm{cov}[X, Y] && (12)
\end{aligned}
$$

Exercise 2.4 Bayes rule for medical diagnosis

$$
X = \begin{cases} 1 & \text{(I have the disease)} \\ 0 & \text{(otherwise)} \end{cases} \tag{13}
$$

$$
Y = \begin{cases} 1 & \text{(testing positive)} \\ 0 & \text{(otherwise)} \end{cases} \tag{14}
$$

$$
\begin{aligned}
P(X=1|Y=1) &= \frac{P(Y=1|X=1) \cdot P(X=1)}{P(Y=1)} && (15)\\
&= \frac{P(Y=1|X=1) \cdot P(X=1)}{P(Y=1|X=1) \cdot P(X=1) + P(Y=1|X=0) \cdot P(X=0)} && (16)\\
&= \frac{\frac{99}{100} \times \frac{1}{10000}}{\frac{99}{100} \times \frac{1}{10000} + \frac{1}{100} \times \frac{9999}{10000}} && (17)\\
&= \frac{99}{99 + 9999} && (18)\\
&= \frac{1}{102} && (19)\\
&\simeq 9.8 \times 10^{-3} && (20)
\end{aligned}
$$

Exercise 2.5 The Mohty Hall Problem

$$X = \begin{cases} 1 & \text{(the prize in door 1)} \\ 2 & \text{(the prize in door 2)} \\ 3 & \text{(the prize in door 3)} \end{cases} \tag{21}$$

$$Y = \begin{cases} 1 & \text{(the host opens door 1)} \\ 2 & \text{(the host opens door 2)} \\ 3 & \text{(the host opens door 3)} \end{cases} \tag{22}$$

$$P(X=1) = P(X=2) = P(X=3) = \frac{1}{3} \tag{23}$$

$$\begin{aligned}
P(Y=3|X=1) &= \frac{1}{2} && (24)\\
P(Y=3|X=2) &= 1 && (25)\\
P(Y=3|X=3) &= 0 && (26)\\
P(X=1|Y=3) &= \frac{P(Y=3|X=1)\cdot P(X=1)}{P(Y=3)} && (27)\\
&= \frac{P(Y=3|X=1)\cdot P(X=1)}{\sum_{i=1}^{3} P(Y=3|X=i)\cdot P(X=i)} && (28)\\
&= \frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{3}+1\times\frac{1}{3}+0\times\frac{1}{3}} && (29)\\
&= \frac{1}{3} && (30)\\
P(X=3|Y=3) &= 0\ (\because P(Y=3|X=3)=0) && (31)\\
P(X=2|Y=3) &= 1 - P(X=1|Y=3) - P(X=3|Y=3) && (32)\\
&= \frac{2}{3} && (33)
\end{aligned}$$

This result shows the contestant should switch to door 2.

Exercise 2.6 Conditional Independence

(a)

$$\begin{aligned}
P(H, e_1, e_2) &= P(e_1, e_2|H)P(H) && (34)\\
P(H, e_1, e_2) &= P(H|e_1, e_2)p(e_1, e_2) && (35)\\
\therefore p(H|e_1, e_2) &= \frac{P(e_1, e_2|H)P(H)}{P(e_1, e_2)} && (36)
\end{aligned}$$

Above calculation shows (ii.) is sufficient.

(b)

$$P(H|e_1,e_2) = \frac{P(e_1,e_2|H)P(H)}{P(e_1,e_2)} \tag{37}$$
$$= \frac{P(e_1|H)P(e_2|H)P(H)}{P(e_1,e_2)} (\because E_1 \text{ and } E_2 \text{ are conditionally independent given H}) \tag{38}$$
$$P(e_1,e_2) = \sum_{i=1}^{K} P(e_1,e_2|H=i)P(H=i) \tag{39}$$
$$= \sum_{i=1}^{K} P(e_1|H=i)P(e_2|H=i)P(H=i) \tag{40}$$

Above calculation shows (i.), (ii.) and (iii.) are sufficient.

Exercise 2.7 Pairwise independence does not imply mutual independence

(Reference to http://www.cut-the-knot.org/Probability/MutuallyIndependentEvents.shtml)

There are four balls numbered as below.

$$B1 = 110$$
$$B2 = 101$$
$$B3 = 011$$
$$B4 = 000$$

For $k = 1, 2, 3$ let A_k be the event of drawing a ball with 1 in the kth position.

$$P(A_1) = \frac{1}{2},\ P(A_2) = \frac{1}{2},\ P(A_3) = \frac{1}{2} \tag{41}$$
$$P(A_1,A_2) = \frac{1}{4},\ P(A_1,A_3) = \frac{1}{4},\ P(A_2,A_3) = \frac{1}{4} \tag{42}$$
$$\therefore P(A_1,A_2) = P(A_1)P(A_2) \tag{43}$$
$$P(A_1,A_3) = P(A_1)P(A_3) \tag{44}$$
$$P(A_2,A_3) = P(A_2)P(A_3) \tag{45}$$

This shows all pairs of variables are pairwise independence.

$$P(A_1,A_2,A_3) = 0 \tag{46}$$
$$P(A_1)P(A_2)P(A_3) = \frac{1}{8} \tag{47}$$
$$\therefore P(A_1,A_2,A_3) = P(A_1)P(A_2)P(A_3) \tag{48}$$

This shows mutual independence doesn't hold.

Exercise 2.8 Conditional independence iff joint factorizes

($\Rightarrow$)

$$g(x,z) = p(x|z), h(y,z) = p(y|z) \tag{49}$$

($\Leftarrow$)

$$p(x,y|z) = g(x,z)h(y,z) \tag{50}$$

We calculate the integral of left side and right side.

$$\int_x p(x,y|z)dx = \int_x g(x,z)h(y,z)dx \tag{51}$$

$$p(y|z) = h(y,z)G(x,z) \tag{52}$$

We calculate the integral of left side and right side.

$$\int_y p(y|z)dy = \int_y h(y,z)G(x,z)dy \tag{53}$$

$$H(y,z)G(x,z) = 1 \tag{54}$$

We calculate the integral of left side and right side of (50).

$$\int_y p(x,y|z)dy = \int_y g(x,z)h(y,z)dy \tag{55}$$

$$p(x|z) = g(x,z)H(y,z) \tag{56}$$

Finally, we get

$$\begin{aligned} p(x,y|z) &= g(x,z)h(y,z) && (57)\\ &= \frac{p(x|z)}{H(y,z)}\frac{p(y|z)}{G(x,z)} && (58)\\ &= p(x|z)p(y|z) && (59) \end{aligned}$$

Exercise 2.9 Conditional independence

(a)

$$
\begin{aligned}
(X \perp W|Z,Y) &\Leftrightarrow p(X,W|Z,Y) = p(X|Z,Y)p(W|Z,Y) && (60)\\
(X \perp Y|Z) &\Leftrightarrow p(X,Y|Z) = p(X|Z)p(Y|Z) && (61)\\
(X \perp Y,W|Z) &\Leftrightarrow p(X,W|Z) = p(X|Z)p(W|Z) && (62)\\
p(X,W,Z|Y) &= p(X,Y,Z)p(W|Z,Y) \ (\because (60)) && (63)\\
p(X,Y,Z) &= p(X,Z)p(Y|Z) \ (\because (61)) && (64)
\end{aligned}
$$

Because of (63) and (64), we get

$$
\begin{aligned}
p(X,Y,Z,W) &= p(X,Z)p(Y|Z)p(W|Z,Y) && (65)\\
&= p(X,Z)p(Y,W|Z) && (66)
\end{aligned}
$$

We calculate the integral of left side and right side.

$$
\begin{aligned}
\int_Y p(X,Y,Z,W)dY &= \int_Y p(X,Z)p(Y,W|Z)dY && (67)\\
p(X,Z,W) &= p(X,Z)p(W|Z) && (68)\\
\frac{p(X,W|Z)}{p(Z)} &= \frac{p(X|Z)}{p(Z)}p(W|Z) && (69)\\
p(X,W|Z) &= p(X|Z)p(W|Z) && (70)
\end{aligned}
$$

We showed (60) $\wedge$ (61) $\Rightarrow$ (62).

(b)

Somebody help me!

Exercise 2.10 Deriving the inverse gamma density

$$
\begin{aligned}
p_y(y) &= p_x(x)|\frac{dx}{dy}| && (71)\\
&= \frac{b^a}{\Gamma(a)}x^{a-1}e^{-xb}|-\frac{1}{y^2}| && (72)\\
&= \frac{b^a}{\Gamma(a)}y^{-(a-1)}e^{-\frac{b}{y}}\frac{1}{y^2} && (73)\\
&= \frac{b^a}{\Gamma(a)}y^{-(a+1)}e^{-\frac{b}{y}} && (74)\\
&= IG(y|a,b) && (75)
\end{aligned}
$$

Exercise 2.11 Normalization constant for a 1D Gaussian

$$
\begin{aligned}
Z^2 &= \int_0^{2\pi} \int_0^{\infty} r \exp(-\frac{r^2}{2\sigma^2}) dr d\theta && (76)\\
&= \int_0^{2\pi} d\theta \int_0^{\infty} r \exp(-\frac{r^2}{2\sigma^2}) dr && (77)\\
&= 2\pi \int_0^{\infty} r \exp(-\frac{r^2}{2\sigma^2}) dr && (78)\\
&= 2\pi \left[-\sigma^2 \exp(-\frac{r^2}{2\sigma^2}) \right]_0^{\infty} && (79)\\
&= 2\pi\sigma^2 && (80)\\
\therefore Z &= \sqrt{2\pi\sigma^2} && (81)
\end{aligned}
$$

Exercise 2.12 Expressing mutual information in terms of entropies

$$
\begin{aligned}
I(X,Y) &= \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} && (82)\\
&= \sum_x \sum_y p(x,y) \log \frac{p(x|y)p(y)}{p(x)p(y)} && (83)\\
&= \sum_x \sum_y p(x,y)(-\log p(x) + \log p(x|y)) && (84)\\
&= -\sum_x \sum_y p(x,y) \log p(x) + \sum_x \sum_y p(x,y) \log p(x,y) && (85)\\
&= -\sum_x p(x) \log p(x) + \sum_x \sum_y p(x,y) \log p(x|y) && (86)\\
&= H(x) + \sum_x \sum_y p(x,y) \log p(x|y) && (87)\\
&= H(X) + \sum_x \sum_y p(y)p(x|y) \log p(x|y) && (88)\\
&= H(X) + \sum_y p(y) \sum_x p(x|y) \log p(x|y) && (89)\\
&= H(X) - \sum_y p(y) H(X|Y=y) && (90)\\
&= H(X) - H(X|Y) && (91)
\end{aligned}
$$

$I(X,Y) = H(Y) - H(X|Y)$ can be shown like as below.

Exercise 2.13 Mutual information for correlated normals

$$
\begin{aligned}
I(X,Y) &= -H(X) + H(X_1) + H(X_2) && (92)\\
&= -\frac{1}{2}\log_2[4\pi^2 e^2\sigma^4(1-\rho^2)] + \frac{1}{2}\log_2[2\pi e\sigma^2] + \frac{1}{2}\log_2[2\pi e\sigma^2] && (93)\\
&= \frac{1}{2}\log_2 \frac{4\pi^2 e^2\sigma^4}{4\pi^2 e^2\sigma^4(1-\rho^2)} && (94)\\
&= \frac{1}{2}\log_2 \frac{1}{1-\rho^2} && (95)\\
&= -\frac{1}{2}\log_2(1-\rho^2) && (96)
\end{aligned}
$$

and we get

$$
\begin{cases}
\rho = 1 : I(X_1, X_2) = \infty & \text{(perfect correlation)}\\
\rho = 0 : I(X_1, X_2) = 0 & \text{(independent)}\\
\rho = -1 : I(X_1, X_2) = \infty & \text{(perfect correlation)}
\end{cases}
\tag{97}
$$