HW1

1.3 ($\star\star$) Suppose that we have three coloured boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities $p(r) = 0.2$, $p(b) = 0.2$, $p(g) = 0.6$, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting an apple? If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box? **(10 points)**

1.6 ($\star$) Show that if two variables x and y are independent, then their covariance is zero. **(10 points)**

1.10 ($\star$) **www** Suppose that the two variables x and z are statistically independent. Show that the mean and variance of their sum satisfies

$$\begin{aligned} \mathbb{E}[x+z] &= \mathbb{E}[x] + \mathbb{E}[z] \qquad (1.128) \\ \mathrm{var}[x+z] &= \mathrm{var}[x] + \mathrm{var}[z]. \qquad (1.129) \end{aligned}$$

(10 points)

1.14 ($\star\star$) Show that an arbitrary square matrix with elements w_{ij} can be written in the form $w_{ij} = w_{ij}^{\mathrm{S}} + w_{ij}^{\mathrm{A}}$ where w_{ij}^{S} and w_{ij}^{A} are symmetric and anti-symmetric matrices, respectively, satisfying $w_{ij}^{\mathrm{S}} = w_{ji}^{\mathrm{S}}$ and $w_{ij}^{\mathrm{A}} = -w_{ji}^{\mathrm{A}}$ for all i and j. Now consider the second order term in a higher order polynomial in D dimensions, given by

$$\sum_{i=1}^{D}\sum_{j=1}^{D} w_{ij}x_i x_j. \qquad (1.131)$$

Show that

$$\sum_{i=1}^{D}\sum_{j=1}^{D} w_{ij}x_i x_j = \sum_{i=1}^{D}\sum_{j=1}^{D} w_{ij}^{\mathrm{S}}x_i x_j \qquad (1.132)$$

so that the contribution from the anti-symmetric matrix vanishes. We therefore see that, without loss of generality, the matrix of coefficients w_{ij} can be chosen to be symmetric, and so not all of the D^2 elements of this matrix can be chosen independently. Show that the number of independent parameters in the matrix w_{ij}^{S} is given by $D(D+1)/2$.

(10 points)

1.28 ($\star$) In Section 1.6, we introduced the idea of entropy $h(x)$ as the information gained on observing the value of a random variable x having distribution $p(x)$. We saw that, for independent variables x and y for which $p(x, y) = p(x)p(y)$, the entropy functions are additive, so that $h(x, y) = h(x) + h(y)$. In this exercise, we derive the relation between h and p in the form of a function $h(p)$. First show that $h(p^2) = 2h(p)$, and hence by induction that $h(p^n) = nh(p)$ where n is a positive integer. Hence show that $h(p^{n/m}) = (n/m)h(p)$ where m is also a positive integer. This implies that $h(p^x) = xh(p)$ where x is a positive rational number, and hence by continuity when it is a positive real number. Finally, show that this implies $h(p)$ must take the form $h(p) \propto \ln p$.

(15 points)

HW1

1.32 ($\star$) Consider a vector $\mathbf{x}$ of continuous variables with distribution $p(\mathbf{x})$ and corresponding entropy $\mathrm{H}[\mathbf{x}]$. Suppose that we make a nonsingular linear transformation of $\mathbf{x}$ to obtain a new variable $\mathbf{y} = \mathbf{A}\mathbf{x}$. Show that the corresponding entropy is given by $\mathrm{H}[\mathbf{y}] = \mathrm{H}[\mathbf{x}] + \ln |\mathbf{A}|$ where $|\mathbf{A}|$ denotes the determinant of $\mathbf{A}$.

(15 points)

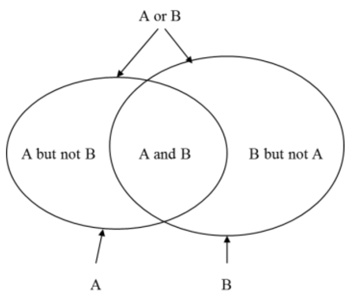
1.39 ($\star\star\star$) Consider two binary variables x and y having the joint distribution given in Table 1.3.

Evaluate the following quantities

(a) $\mathrm{H}[x]$	**(c)** $\mathrm{H}[y\|x]$	**(e)** $\mathrm{H}[x, y]$
(b) $\mathrm{H}[y]$	**(d)** $\mathrm{H}[x\|y]$	**(f)** $\mathrm{I}[x, y]$.

Draw a diagram to show the relationship between these various quantities.

A diagram example



(10 points)

8. (20 points) Let $f(x) = x\left[C - \log\left(\sum_{n=1}^{N} e^{\frac{-d_n}{x}}\right)\right]$ defined over $x > 0$, where C and $\{d_n\}_{n=1\ldots N}$ are constants.

(a) Derive $\frac{\partial f(x)}{\partial x}$.

(b) Derive $\frac{\partial^2 f(x)}{\partial x^2}$.

(c) Is $f(x)$ a convex function?