

Exercise 2: Introduction to Probability – Solutions –

Exercise 2.1: Conditional Probability

a) Probability of drawing an apple:

$$\begin{aligned}
 p(a) &= \sum_{box} p(a, box) \\
 &= \sum_{box} p(a|box)p(box) \\
 &= p(a|r)p(r) + p(a|b)p(b) + p(a|g)p(g) \\
 &= 0.3 * 0.2 + 0.5 * 0.2 + 0.3 * 0.6 = 0.34
 \end{aligned}$$

b) Probability of green box given orange

$$\begin{aligned}
 p(g|o) &= \frac{p(g,o)}{p(o)} \\
 &= \frac{p(o|g)p(g)}{\sum_{box} p(o|box)p(box)} \\
 &= \frac{0.18}{0.36} = 0.5
 \end{aligned}$$

Exercise 2.2: Bayes' Rule

a) Applying the Bayes' theorem:

$$p(x_i|z) = \frac{p(z|x_i)p(x_i)}{\sum_{i=1}^3 p(z|x_i)p(x_i)} \quad (1)$$

$$p(x_1|z) = \frac{0.8*1/3}{0.8*1/3+0.4*1/3+0.1*1/3} = 0.616 \quad (2)$$

$$p(x_2|z) = \frac{0.4*1/3}{0.8*1/3+0.4*1/3+0.1*1/3} = 0.308 \quad (3)$$

$$p(x_3|z) = \frac{0.1*1/3}{0.8*1/3+0.4*1/3+0.1*1/3} = 0.077 \quad (4)$$

Exercise 2.3: Expectation, Variance and Covariance

a) Since x and z are independent, their joint distribution factorizes $p(x, z) = p(x)p(z)$:

$$\begin{aligned} E[x + z] &= \int \int (x + z)p(x)p(z)dx dz \\ &= \int xp(x)dx + \int zp(z)dz \\ &= E[x] + E[z] \end{aligned} \tag{5}$$

Similarly for the variances, we first note that

$$(x + z - E[x + z])^2 = (x - E[x])^2 + (z - E[z])^2 + 2(x - E[x])(z - E[z]) \tag{6}$$

where the final term will integrate to zero with respect to the factorized distribution $p(x)p(z)$. Hence:

$$\begin{aligned} \text{var}[x + z] &= \int \int (x + z - E[x + z])^2 p(x)p(z)dx dz \\ &= \int (x - E[x])^2 p(x)dx + \int (z - E[z])^2 p(z)dz \\ &= \text{var}[x] + \text{var}[z] \end{aligned} \tag{7}$$

b)

$$\text{cov}[xz] = E[xz] - E[x]E[z]$$

If x and y are independent:

$$\begin{aligned} E[xz] &= \sum_x \sum_y xyp(x, y) \\ &= \sum_x \sum_y xyp(x)p(y) \\ &= \sum_x xp(x) \sum_y yp(y) \\ &= E[x]E[y] \end{aligned} \tag{8}$$

c) The covariance matrix is not valid because of the following reasons;

- The matrix is not positive semi-definite, easily checked by
- The eigenvectors are $\begin{bmatrix} -0.6569 \\ 10.6569 \end{bmatrix}$; should be all real and positive. Use **eig** in Matlab to check this.
- Determinant is -4 ; should be positive as a consequence of the eigenvalues. Use **det** in Matlab

Exercise 2.4: Probability distributions and moments

See provided Matlab files! First load the data, then run the scripts `CalcStats` followed by `PlotData`.