BIAS VARIANCE PROBLEM SHEET

When modelling systems with uncertainty it is convenient to define *random variables*. The are numbers that we associate with the outcome of some stochastic event. We associate a probability (or probability density) with the set out outcomes such that the random variables take a particular value. We often write random variables using capital letters (e.g. X) while the actual values the X takes we write with small letters x. Thus $\mathbb{P}(X = x)$ is the probability that the random value, X, takes value, x. We write non-random variables (scalars) with a small letter (e.g. c). Note that for most continuous random variables $\mathbb{P}(X = x) = 0$ so instead we define the probability density

$$f_X(x) = \lim_{\delta x \to 0} \frac{\mathbb{P}(x \leq X \leq x + \delta x)}{\delta x}.$$

The expectation (or average value) of some function, g, of X is written as

$$\mathbb{E}_X\left[g(X)\right] = \begin{cases} \sum_{x \in \mathcal{X}} \mathbb{P}(X = x)\, g(x) \\ \int f_X(x)\, g(x)\, \mathrm{d}x \end{cases}$$

depending on whether X is a continuous or discrete random variable. Note $\mathcal{X}$ is the possible values that the random variable can take. When it is clear what random variables we are taking expectation with respect to then we will often write $\mathbb{E}\left[\cdot\right]$ for $\mathbb{E}_X\left[\cdot\right]$.

1

(a) Let X be outcome of an honest dice ($\mathcal{X} = \{1,2,3,4,5,6\}$). What is

(i) $\mathbb{E}\left[X\right]$

(ii) $\mathbb{E}\left[2X\right]$

(iii) $\mathbb{E}\left[X^2\right]$

[3 marks]

(i) $\mathbb{E}\left[X\right] = \frac{1+2+3+4+5+6}{6} = 3.5$

(ii) $\mathbb{E}\left[2X\right] = 2\mathbb{E}\left[X\right] = 7$

(iii) $\mathbb{E}\left[X^2\right] = \frac{1+4+9+16+25+36}{6} = \frac{91}{6} = 15.1666$

(b) Let X be a random variable as before and Y be a random variable for a second independent dice. What is

(i) $\mathbb{E}_X\left[X + Y\right]$

(ii) $\mathbb{E}_{X,Y}\left[X + Y\right]$

(iii) $\mathbb{E}_X\left[XY\right]$

(iv) $\mathbb{E}_{X,Y}\left[XY\right]$

[4 marks]

(i) $\mathbb{E}_X\left[X+Y\right] = \mathbb{E}_X\left[X\right] + Y = 3.5 + Y$

(ii) $\mathbb{E}_{X,Y}\left[X+Y\right] = \mathbb{E}_X\left[X\right] + \mathbb{E}_Y\left[Y\right] = 7$

(iii) $\mathbb{E}_X\left[XY\right] = \mathbb{E}_X\left[X\right] Y = 3.5Y$

(iv) $\mathbb{E}_{X,Y}\left[XY\right] = \mathbb{E}_X\left[X\right] \mathbb{E}_Y\left[Y\right] = 3.5^2 = 12.25$

(c) Let X be the random variable as before and E be a random variable equal to 0 if X is odd and 1 if X is even. Note that E is not independent of X. What is

(i) $\mathbb{P}(E=1) (= \mathbb{E}_X\left[E\right])$

(ii) $\mathbb{E}_X\left[X+E\right]$

(iii) $\mathbb{E}_X\left[XE\right]$

[3 marks]

(i) $\mathbb{P}(E=1) = \mathbb{E}_X\left[E\right] = \frac{0+1+0+1+0+1}{6} = 0.5$

(ii) $\mathbb{E}_X\left[X+E\right] = \mathbb{E}_X\left[X\right] + \mathbb{E}_X\left[E\right] = 3.5 + 0.5 = 4$

(iii) $\mathbb{E}_X\left[XE\right] = \frac{0+2+0+4+0+6}{6} = 2$

End of question 1

Note that expectations are *linear operators* so that

$$\mathbb{E}_x\left[a g(X) + b h(X)\right] = a\mathbb{E}_X\left[g(X)\right] + b\mathbb{E}_x\left[h(X)\right].$$

Note that this also means

$$\mathbb{E}_X\left[\sum_{i=1}^{n} g_i(X)\right] = \sum_{i=1}^{n} \mathbb{E}_X\left[g_i(X)\right].$$

If X and Y are independent (so $\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\mathbb{P}(Y=y)$) then

$$\mathbb{E}_{X,Y}\left[g(X)h(Y)\right] = \mathbb{E}_X\left[g(X)\right]\mathbb{E}_Y\left[h[Y]\right].$$

However is X and Y are not independent random variables then

$$\mathbb{E}_{X,Y}\left[g(X)h(Y)\right] = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathbb{P}(X=x, Y=y)\, g(X)h(Y)$$

and usually $\mathbb{E}_{X,Y}\left[g(X)h(Y)\right] \neq \mathbb{E}_X\left[g(X)\right]\mathbb{E}_Y\left[h[Y]\right]$.

END OF PAPER