MATH 7501: Probability and Statistics

Isidoros Strouthos University College London 2017-2018

Exercise sheet 1 (completion)

Exercise 4

In part (a) of this exercise, we are first asked to find the forecast value $\hat{p}$ that maximises the expected forecast score, given that the true probability of rain is p.

In order to determine this, we start by determining an expression for the expected forecast score, $E(S)$, by using the "formula" for S itself, i.e. by using $S = Xln\hat{p} + (1-X)ln(1-\hat{p})$. Using this, and some of the basic properties of expectation from your lectures, we obtain:

$$\begin{aligned} E(S) &= E\left(Xln\hat{p} + (1-X)ln(1-\hat{p})\right) \\ &= E\left(Xln\hat{p}\right) + E\left((1-X)ln(1-\hat{p})\right) \\ &= E(X)ln\hat{p} + E(1-X)ln(1-\hat{p}) \\ &= E(X)ln\hat{p} + (1-E(X))\,ln(1-\hat{p}) \end{aligned}$$

The above shows that we may find an expression for $E(S)$ by using the value of $E(X)$. This is a value that is essentially provided to us; it is stated that the true probability of rain is p. As a result, X (which is said to take the value 1 if it rains and 0 otherwise), takes the value 1 with probability p and the value 0 with probability $1-p$. So:

$$E(X) = 1 \times p + 0 \times (1-p) = p$$

Substituting this in the resulting expression for $E(S)$, from above, gives:

$$E(S) = pln\hat{p} + (1-p)ln(1-\hat{p})$$

We are now asked to maximise this expected score with respect to $\hat{p}$ and we can achieve this by considering the (partial) derivative of $E(S)$ with respect to $\hat{p}$ (it might be helpful to note that the probability p, given in the exercise, is a constant as far as this derivative is concerned):

$$\frac{\partial E(S)}{\partial \hat{p}} = \frac{p}{\hat{p}} - \frac{1-p}{1-\hat{p}} \qquad \left(\text{using } \frac{\partial\left(ln\hat{p}\right)}{\partial \hat{p}} = \frac{1}{\hat{p}}\,,\ \frac{\partial\left(ln(1-\hat{p})\right)}{\partial \hat{p}} = \frac{-1}{1-\hat{p}}\right)$$

We may then set this partial derivative to be equal to zero, in order to obtain

$$\frac{p}{\hat{p}} - \frac{1-p}{1-\hat{p}} = 0 \qquad \text{i.e.} \qquad \frac{p}{\hat{p}} = \frac{1-p}{1-\hat{p}}$$

Rearranging this leads to $p(1-\hat{p}) = (1-p)\hat{p}$, i.e. to $p - p\hat{p} = \hat{p} - p\hat{p}$, and, therefore, to:

$$\hat{p} = p$$

We may then confirm that this value *maximises* the expected score, perhaps by showing that the second partial derivative of $E(S)$ with respect to $\hat{p}$ is negative if $\hat{p} = p$, or by noting that:

- The function $E(S)$ is a continuous function of $\hat{p}$, for values of $\hat{p}$ between 0 and 1.
- The first derivative of the function $E(S)$, with respect to $\hat{p}$, satisfies:

$$\frac{\partial E(S)}{\partial \hat{p}} \rightarrow +\infty \text{ as } \hat{p} \rightarrow 0 \qquad \text{and} \qquad \frac{\partial E(S)}{\partial \hat{p}} \rightarrow -\infty \text{ as } \hat{p} \rightarrow 1$$

Together, these two features indicate that the single turning point for the function, at $\hat{p} = p$, is, indeed, a maximum.

(You might find it helpful to visualise this, e.g. by thinking of the graph of $y = -x^2$, which also involves a continuous function with gradient tending to 'plus infinity' to the left and tending to 'minus infinity' to the right; this, like the function that we are considering here, has a single turning point that is a maximum.)

In any case, now that we have determined that the expected forecast score is maximised if $\hat{p}$ takes the value p, we can solve the final part of (a).

The expected forecast score, $E(S)$, is an idealised long-run average score. So, if forecasters wish to maximise their bonuses/salaries in the long run, they should aim to forecast in a way that would maximise this expected score.

Since $E(S)$ is maximised when $\hat{p}$ (forecast probability of rain) takes the value p (true probability of rain), the forecasters should aim to make "honest" forecasts, i.e. to forecast the true probabilities of rain, in order to maximise their salaries in the long run.

Let us now turn our attention to part (b) of this exercise; here, we are essentially provided with two possible values of p (based on observations from the last 50 years), namely $p = 0.69$ for the West of Ireland, and $p = 0.24$ for New South Wales, Australia.

It turns out that, if forecasters do make "honest" forecasts, as described above, the expected values of the forecast scores in these two cases are similar. In particular, if we set $\hat{p} = p$ in the expression for $E(S)$ we obtain the expected "honest forecast score" of

$$p\, lnp \,+\, (1-p)\, ln(1-p)$$

Substituting $p = 0.69$ into this expression leads to an expected score of approximately -0.619, while, substituting $p = 0.24$ leads to an expected score of approximately -0.551.

In a sense, these values may be considered "close enough" to allow us to use the given scoring rule in order to informally compare the performance of forecasters in the two areas, if we assume that there is little variation in the true day-to-day probabilities of rain for each area.

For example, informal comparison might be possible if, for each of the two regions, the probability of rain on any given day is the same as the probability of rain on any other day.

However, any such comparison might not be appropriate if there is significant variation in the true day-to-day probabilities of rain for each area.

For example, let us consider an "idealised" scenario in which, for each of two regions, rain has occurred on around 25% of all days in the last 50 years, but where, in one region, it is as likely to rain on any one day as on any other day, while in the other, it is highly likely that it will rain on any day in a particular three-month period in a year, and highly unlikely that it will rain on any day in the remaining nine-month period in a year (i.e. in the second case, there is a significant day-to-day variation in the true probability of rain, unlike in the first case). In such a scenario, it might be inappropriate to compare the performances of forecasters in the two regions by using the scoring rule studied in this exercise.