CAB203 Assessment 3 Probability:

The probability problem looks at finding the best crop insurance to buy for the year over a 20-year period. There are 6 fields each with their own failure probabilities, it is unknown which field is Charlies’ and as such which should be used to determine the maximum profit. A python function will need to be developed that takes in the premiums (), farming costs for the year (), money received when there is no crop failure (), the previous year’s crop outcome () and state () representing all the information across each year. This will be used to determine the best crop insurance by using a formula that iterates over the probability data to determine which is the correct field based on precious iterations, the , and then using the newly determined probability data update this probability data for the next iteration, .

To solve the follow problem a formula needs to be developed that follows the listed properties:

1. The probability of failure for each type should be calculated based on the given field data over 20 years.
2. Determine the field type probability using the formula’s outcome of the previous years. None if it’s the first time. Each iteration will use failure type probabilities to help refine the decision making this is also known as the .
3. The calculated value of each insurance should be produced given the probabilities of the calculated different crop failure types.
4. The values need to be compared to determine which insurance will provide the most amount for profit for that year.
5. Based on the insurance chosen for that year use the probability data of the failure types for the next year to further improve in determining the field type the next year, this is the .

The probabilities of all events over the 20-year period including all crop failure types need to be defined. The probability of event formula can be used to determine the probability of each failure type:

For the first year and iteration of the probability data the approximate average of failure events is calculated across all fields. From the field failure data given by Charlie’s book and using the probability event formula, the field failure probability variables can be defined as the following:

* Let be the probability of a drought.
* Let be the probability of hail.
* Let be the probability of grasshoppers.
* Let be the probability of no failure.

The Calculated profit gain for each insurance type based on each failure type can be determined by letting be the insurance payout be the calculated profit for each failure type:

* Comprehensive Insurance:
* Hail Insurance:
* Grasshopper Insurance:
* Basic Insurance:

(NOTE MAYBE REWRITE CALCULATION EQUATIONS IN DIFFERRENT NOTATION)

Let be , let be and let be . The Payout for calculating the payout for each insurance can be calculated as:

The formula for calculating the profit for comprehensive, hail and grasshopper insurance can be written as:

The formula for calculating the profit for basic insurance can be written as:

Using Bayesian’s rule, the belief of Charlies field given the probability of field failure types can be solved:

is the probability of having the field given previous years probability, the . is the previous year’s probability outcome, the . is likelihood of failure given the previous year’s outcome. is the updated belief of the total probability of the failure for the field. The state, , is used to determine the likelihood of which field Charlie owns, this is done by altering the probabilities of failure types, , for the given fields. For the first year the state is set to which means ) won’t be altered. All fields will be assumed to have the same probability out of all possible fields. Let be each possible field:

The following formulas represents our probability of each field, with following probability being the initial belief the probability of the field being Charlies.

The conditional probability formula can be used when determining the likelihood of failure type for a given field:

The given failure type given by helps determines the likelihood as represented by . The likelihood of each field given the failure type can be defined as:

Given the joint probability formula:

As well as the marginal likelihood formula:

Using joint probability, the product of the and the likelihood can be used with normative decision theory to determine the altered probability of each failure type across all fields:

Using the posterior probability formula:

A formula can be created that represents updating the belief of Charlies field given a failure type:

These formulas allow so the belief of the probability of each field and their probability of failure can continually be updated ensuring improved decision making for each year or iteration. With this the probability of failure type is altered as our belief of which field is Charlies is improved.

Using the normative decision theory formula:

The best insurance, , can be determined by:

Using this formula the profit for each insurance can be compared to determine the optimal insurance policy that will provide the maximum profit for that year.

Python code explanation:

Notes:

Can be used for the probability of each field.

Combine the state information of each year into state, , so it contains all prior year data.

Joint distribution formula:

Can be used to find the probability of field failure given the failure. Could also be used for insurance.

* Best insurance to buy each year to maximise profits over 20 years
  + Comprehensive. Pays 80% of your contract price in the event of any kind of crop failure.
  + Hail. Pays 80% of your contract price in the event of crop failure due to hail.
  + Grasshopper. Pays 80% of your contract price in the event of crop failure due to grasshoppers
  + Basic. Pays 50% of your contract price in the event of crop failure not due to hail.
* parameters:
  + Premiums: dictionary of insurance plans with names of insurance being the key.
  + inputCost: fixed number of the cost for farming for the year.
  + contractPrice: a number for the amount of money received if there are no crop failures.
  + lastYearOutcome: a string containing: 'drought', 'hail', 'grasshoppers', 'no failure'. First iteration when there is no previous year it will be None.
  + state: set to None for the first year it will be set to the state variable the that the chooseCropInsurance returns. Used to remember information year to year.
* Should return (insurance, state), insurance is a string choosing one of 'comprehensive', 'hail', 'grasshopper', or 'basic'. State is used as previously mentioned to track data
* Function will be executed 20 times to represent 20 years while being run 5000 times each time selecting a random field for each year. The average of this will be the final total
* Function will be tested with P − B / 1830 P is profit and B = 61000 which is the average net profit of the basic statergy
* Function in probability that could be used:
  + Prob
  + probEvent, calculating probability of each event
  + isProbDist,
  + conditionalProbDistribution, calculating the loss of each insurance type
  + conditionalProb
  + marginalLikelihood,
  + posterior
  + utility
  + decide, choosing insurance