QUAN 102 Statistics for Business

Tutorial 4

Natt Krongkajonsook

4.1 Probabilities

- Probability quantifies how certain we are about future events.
- Examples of probabilities
  - There's a 70% chance of rain this evening.
  - With probability 0.6 Apple will close up for the day.
  - The degree of my belief in passing the exam is 0.9.

4.2 Notations

Suppose A and B are two events of interest.

- $P(A)$ is the probability that A will happen. $P(B)$ is the probability that B will happen.
- $P(A \cap B)$ is the probability that both A and B will happen.
- $P(A \cup B)$ is the probability that A or B will happen. In other words, $P(A \cup B)$ is the probability that either A or B (or both) will occur.
- $P(A|B)$ is the probability that A will occur given that B has occurred already.
- $P(A')$ is the probability that A will not happen. $P(B')$ is the probability that B will not happen.

4.3 Theoretical VS Empirical Probability

- Theoretical probability (based on facts)
  - The probability of an event occurring is number of occurrences of that event divided by total number of possible outcomes.
  - Example: coin tossing
    * Only can come up heads once.
    * Two possibilities: head or tail.
    * Probability of a fair coin coming up heads $= \frac{1}{2}$.
- Empirical Probability (based on trials)
  - The probability of an event occurring is number of times an event has occurred divided all possible times it could have occurred.
  - Example: coin tossing
    * The coin comes up heads 4 out of 10 times.
    * Probability of a coin coming up heads $= \frac{4}{10} = 0.4$.

4.4 Objective VS Subjective Probability

- An objective probability of an event occurring is calculated by carrying out a certain number of trials and then by seeing how many times such event has occurred.
- A subjective probability of an event occurring is the likelihood of such event occurring based on an individual's personal judgement. It is not based on any precise computation but it is often a reasonable assessment by a knowledgeable person.

4.5 Probability Rules

1. All probabilities are between 0 and 1 inclusive.
2. The sum of all the probabilities in the set of all possible outcomes is 1.
3. The probability of an event which cannot occur is 0.
4. The probability of an event which must occur is 1.
5. $P(A') = 1 - P(A)$
6. In general, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
7. A and B are mutually exclusive if $P(A \cap B) = 0$. In this case, $P(A \cup B) = P(A) + P(B)$.
8. In general, $P(A \cap B) = P(A) \times P(B|A)$. Thus $P(B|A) = \frac{P(A \cap B)}{P(A)}$.
9. Two events are independent if the occurrence of one does not change the probability of the other occurring (Exercise 5.1).
   (a) If A and B are independent, then $P(A|B) = P(A)$, and $P(B|A) = P(B)$.
   (b) As a consequence of 9a, $P(A \cap B) = P(A) \times P(B)$.

4.6 Probability Trees

- Drawing probability trees is a useful way of simplifying many probability problems, especially in situations where actions are taken or decisions are made sequentially.
- Constructing the tree for a sequential process, e.g. the situation in Exercise 5.4:
  - The root of the tree corresponds to the starting point of the process. In this case, the starting point represents ethnicity of defendants.
  - Line segments called "branches" connect the root to "nodes" representing the different outcomes that are possible at the first state of the process. In this case, they are Māori (`M`), Pacific Island (`P`), and European (`E`).
  - Each branch of stage 1 nodes is connected to nodes representing the possible outcomes at the next stage. In this case, bail being granted (`Y`), or bail not being granted (`N`) are nodes of the next stage.
  - If there are more than two stages, this process will continue until all stages are completed.

- The tree in this situation should look like this:

```
 Y N    Y N    Y N
 | |    | |    | |
 | |    | |    | |
  M      P      E
  |      |      |
  |      |      |
```

- Using the above method to draw the tree for the situation in Exercise 5.15, we should get:

```
p s t   p s t   p s t
| | |   | | |   | | |
| | |   | | |   | | |
  P       S       T
  |       |       |
  |       |       |
```

where in this situation, we define `P` = {elder has primary education only}, `S` = {elder has secondary education but no tertiary}, `T` = {elder has tertiary education}, and `p`, `s`, and `t` are similarly for the younger brother.

- Using probability trees to determine possible outcomes
  - Each path from the starting point to a node at the other end of the tree represents one outcome of the complete process.
  - Thus, in Exercise 5.4, there are six outcomes: `MY`, `MN`, `PY`, `PN`, `EY`, and `EN`
  - Similarly, in Exercise 5.15, there are nine outcomes: `Pp`, `Ps`, `Pt`, `Sp`, `Ss`, `St`, `Tp`, `Ts`, and `Tt`.

- Labelling the branches with conditional probabilities
  - **Branch probabilities at stage 1**: Place the probabilities of the stage 1 outcomes on the branches that lead from the starting point to the outcome:
    * In Exercise 5.4, of all defendants, 35% were `M`, 15% were `P`, and 50% were `E`.
    * In Exercise 5.15, of all elder brothers, 5% were `P`, 80% were `S`, and 15% were `T`.
  - **Branch probabilities at all other stages**: Place *conditional* probabilities on the branches at later stages. The condition is the history of choices represented by the path from the start to the beginning of the branch.
    * In Exercise 5.4, conditional on `M`, the probability of `Y` is 40%, and the probability of `N` is 60%. Similarly the probability of `Y` given `P` is 40%, the probability of `N` given `P` is 60%, the probability of `Y` given `E` is 80%, and the probability of `N` given `E` is 20%.
    * In Exercise 5.15, the probability of `p` given `P` is 80%, and so on.

- Using the tree to calculate probabilities
  - **Probabilities of individual outcomes**: To determine the probability of an outcome, multiply the probabilities along its path. This is a consequence of Rule 8. For example:
    - In Exercise 5.4, the probability of `MY` is a product of the probability of `M` and the probability of `Y` given `M`, which is $.35 \times .4 = .14$. Similar criteria can be used to determine probabilities of `MN`, `PY`, `PN`, `EY`, and `EN`.
    - In Exercise 5.15, the probability of `Pp` is a product of the probability of `P` and the probability of `p` given `P`, which is $.05 \times .8 = .04$. Similar criteria can be used to determine probabilities of `Ps`, `Pt`, `Sp`, `Ss`, `St`, `Tp`, `Ts`, and `Tt`.
  - **Probabilities of compound events**: Since different paths in the tree diagram represent mutually exclusive events, we can add their probabilities without worrying about any overlap to find the probability of a compound event. This is a consequence of Rule 7. For example:
    - In Exercise 5.4, the probability of `Y` is a sum of probabilities of `MY`, `PY`, and `EY`, which is $.35 \times .4 + .15 \times .06 + .5 \times .8 = .4$.
    - In Exercise 5.15, the probability that a pair of brothers will have the same level of education is a sum of probabilities of `Pp`, `Ss`, and `Tt`, which is $.05 \times .8 + .8 \times .5 + .15 \times .4 = .06$.
- Now use appropriate probability rules and the tree diagram to finish Exercise 5.15 and 5.17.

4.7 Probability Distributions

Probability distribution is a list of all possible outcomes and their associated probabilities which can occur within a particular population of interest. For example:

- In Exercise 5.4, the probability distribution should look like this

Outcomes	MY	MN	PY	PN	EY	EN
Probabilities	.14	.21	.06	.09	.4	.1

- In Exercise 5.15, the probability distribution should look like this

Outcomes	Pp	Ps	Pt	Sp	Ss	St	Tp	Ts	Tt
Probabilities	.04	.01	0	.08	.40	.32	0	.09	.06