STOCHASTIC MODELLING

STAT6102
Fall 2018, CUHK

Yuanyuan Lin (Instructor)

Lecture Notes 1

Preliminarily Concepts and Foundations

We review some basic concepts and foundations about probability, in preparation for the formal contents of this course.

Example 1*. De Méré's Problem. (PROBABILITY SPACE)

The Chevalier de Méré was a French nobleman and a gambler of the 17th century. He was interested in two gambling games: A: betting on at least one ace turns up with four rolls of a die; B: betting at least one double-ace turns up in 24 rolls of two dice. It is tempting to believe the two events have same chance to occur. His reasoning:

In one roll of a die, there is 1/6 chance of getting an ace, so in 4 rolls, there is $4*(1/6) = 2/3$ chance of getting at least one die.

In one roll of two dice there is 1/36 chance of getting a double-ace, so in 24 rolls, there is $24 * (1/36) = 2/3$ chance of getting at least one double ace.

Therefore the two events, by this reasoning, have same chance. However, the chance computed here, 2/3 obviously contracted his gambling experience that the chances are somewhat close to half, making the betting games a nearly fair game.

De Méré turned to Blaise Pascal (1623-1662) and Pierre de Fermat (1601-1665) for help. And the two mathematicians/physicists pointed out that the above reasoning is erroneous and gave the right answer:

$P(\text{getting at least one ace in 4 rolls of a die}) = 1 - (1 - 1/6)^4 = 0.518$

4-rolls makes favorable bet, and 3-rolls makes not.

$P(\text{getting at least one double-ace in 24 rolls of two dice}) = 1 - (1 - 1/36)^{24} = 0.491$

25 rolls makes a favorable bet, 24 rolls make still an unfavorable bet.

Historical remark: (Cardano's mistake) Probability theory has an infamous birth place: gambling room. The earliest publications on probability dates back to "Liber de Ludo Aleae" (the book on games of chance) by Gerolamo Cardano (1501-1576), an Italian mathematician/physician/ astrologer/gambler of the time. Cardano made an important discovery of the product law of independent events, and it is believed that "Probability" was first coined and used by Cardano. Even though he made several serious mistakes that seem to be elementary nowadays, he is considered as a pioneer who first systematically computed probability of events. Pascal, Fermat, Bernoulli and de Moivre are among many other prominent developers of the probability theory. The axioms of probability was first formally and rigorously shown by Kolmogorov in the last century.

1. Some notations/definitions:

   (a) S - sample space

(b) E, F - events

(c) $E \cup F$ - E union F

(d) EF or $E \cap F$ - E intersect F

(e) If $EF = \emptyset$, then we say E and F are mutually exclusive.

(f) $\cup_{n=1}^{\infty} E_n$ means $E_1 \cup E_2 \cup E_3 \cup \ldots$

(g) $\prod_{n=1}^{\infty} E_n$ means $E_1 E_2 E_3 \ldots$

(h) E^c - complement of E

2. Probabilities Defined on Events:

Consider an experiment whose sample space is S. For each event E of the sample space S, we assume that a number $P(E)$ is defined and satisfies the following three conditions:
1. $0 \ \leq P(E) \ \leq 1$.
2. $P(S) = 1$.
3. For any sequence of events $E_1, \ E_2, \ldots$ which are mutually exclusive, (i.e. whenever $n \ \neq \ m$, we have $E_n E_m = \emptyset$) then

$$P(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(a)$$

We refer $P(E)$ as the probability of the event E.

Note:

(a) $P(E) + P(E^c) = 1$.

(b) $P(E \cup F) = P(E) + P(F) - P(EF)$.

(c) If E and F are mutually exclusive, then $P(E \cup F) = P(E) + P(F)$.

Example: A urn contains 3 red, 4 green, 5 blue balls, randomly select 2 balls out of the urn. Find the probability of getting (i) 2 red balls, (ii) 1 red 1 green ball, and (iii) at least 1 red ball

Solution: Let R = red ball, G = green ball.

$$(i)\quad P(getting\ 2R) = \frac{C_2^3}{C_2^{12}} = \frac{3}{(12 \times 11/2)} = \frac{1}{22}.$$

$$(ii)\quad P(getting\ 1R\ 1G) = \frac{C_1^3 C_1^4}{C_2^{12}} = \frac{3 \times 4}{(12 \times 11/2)} = \frac{2}{11}.$$

$$(iii)\quad P(at\ least\ 1R) = 1 - P(getting\ no\ red) = 1 - \frac{C_0^3 C_2^9}{C_2^{12}} = 1 - \frac{(9 \times 8)/2}{(12 \times 11/2)} = \frac{1}{6}.$$

3. Conditional Probabilities

$P(E|F)$ denotes the conditional probability that E occurs given that F has occurred. Note that

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}$$

for $P(F) > 0$.

Example: A family has two children. What is the conditional probability that both are boys given that at least one of them is a boy?

Solution: Let E be the event that both children are boys, and let F be the event that at least one of them is a boy. Note that

$$P(EF) = P(E) = \frac{1}{4}; \quad P(F) = 1 - P(getting\ two\ girls) = 1 - \frac{1}{4} = \frac{3}{4}$$

Then

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

4. Independent Events

Two events E and F are said to be independent if

$$P(EF) = P(E)P(F)$$

This implies that E and F are independent if

$$P(E|F) = P(E)$$

This also implies that $P(F|E) = P(F)$.

Two events E and F which are not independent are said to be dependent. The definition of independence can be extended to more than two events. The events $E_1, E_2, \ldots, E_n$ are said to be independent if for every subset $E_{1'}, E_{2'}, \ldots, E_{r'},\ r \le n$,

$$P(E_{1'}E_{2'}\ldots E_{r'}) = P(E_{1'})P(E_{2'})\ldots P(E_{r'}).$$

Example: Suppose we toss two fair dice. Let E be the event that the sum of the dice is six and F be the event that the first die equals 4. Show that E and F are dependent.

Solution:

$$P(EF) = P(\{4,2\}) = \frac{1}{36}; \quad P(E) = \frac{5}{36}; \quad P(F) = \frac{1}{6}.$$
$$P(EF) \neq P(E)P(F).$$

Therefore, E and F are dependent.

Example: Let A and B be independent events with $P(A) = 1/4$ and $P(B) = 2/3$. Find (a) $P(A \cap B)$; (b) $P(A \cap B')$.

$P(A) = \frac{1}{4}$, $P(B) = \frac{2}{3}$

a) $P(A \cap B) = \frac{1}{4}\frac{2}{3} = \frac{1}{6}$ (A and B are independent)

Note that
$P(A \cap B) = P(A)P(B) = P(A)[1 - P(B')] = P(A) - P(A)P(B') = P(A \cap B) + P(A \cap B') - P(A)P(B')$.

Therefore $P(A \cap B') = P(A)P(B') = \frac{1}{4}\frac{1}{3} = \frac{1}{12}$.

5. Bayes' Formula

Suppose that $F_1, F_2, \ldots, F_n$ are mutually exclusive events such that $\cup_{i=1}^{n} F_i = S$, then we have

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}$$

Proof: Note that

$$E = \cup_{i=1}^{n} EF_i$$

and using the fact that the events EF_i, $i = 1, \ldots, n$ are mutually exclusive, we have

$$P(E) = \sum_{i=1}^{n} P(EF_i) = \sum_{i=1}^{n} P(E|F_i)P(F_i).$$

Thus

$$P(F_j|E) = \frac{P(EF_j)}{P(E)} = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}$$

Example: In answering a question on a multiple choice test a student either knows the answer or he guesses. Let p be the probability that he knows the answer and $1 - p$ be the probability that he guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives.
Question: What is the probability that a student knows the answer to a question given that he answered it correctly?

Solution: Let C be the event that the student answers the question correctly and K be the event that he actually knows the answer. Now

$$\begin{aligned} P(K|C) &= \frac{P(KC)}{P(C)} \\ &= \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K^c)P(K^c)} \\ &= \frac{p}{p + (1/m)(1-p)} \\ &= \frac{mp}{1 + (m-1)p}. \end{aligned}$$

Example: A blood disease is present in

2% of the population in a serious form,
10% of the population in a light form,
88% not present.

An imperfect clinical test is 90% successful in detecting the disease in serious cases. [i.e. if a person has a serious case of the disease, the probability of 0.9 that the clinical test will be +ve.]

Moreover, the success rate of the clinical test is 0.6 among the light cases. Among unaffected persons, the probability that the test will be +ve is 0.1.

Question: A person is selected at random from the population is given the test and the result is +ve. What is the probability that this person has a serious case of the disease?

Solution: Denote Serious case: S; Light case: L; No disease: N; Test positive: +.

$$P(S) = 0.02;\ P(L) = 0.1;\ P(N) = 0.88$$

$$P(+|S) = 0.9;\ P(+|L) = 0.6;\ P(+|N) = 0.1$$

$$
\begin{aligned}
P(+) &= P(+\cap S) + P(+\cap L) + P(+\cap N)\\
&= P(S)P(+|S) + P(L)P(+|L) + P(N)P(+|N)\\
&= (0.02)(0.9) + (0.1)(0.6) + (0.88)(0.1) = 0.166.
\end{aligned}
$$

$$
\begin{aligned}
P(S|+) &= \frac{P(+/capS)}{P(+)}\\
&= \frac{P(S)P(+|S)}{P(+)}\\
&= \frac{(0.02)(0.9)}{0.166} = 0.11.
\end{aligned}
$$

Example 2*. Jailer's reasoning. (Bayes formula)

Three men, A, B and C are in jail and one to be executed and the other two to be freed. C, being anxious, asked the jailer to tell him who of A and B would be freed. The jailer, pondering for a while, answered *"for your own interest, I will not tell you, because, if I do, your chance of being executed would rise from 1/3 to 1/2."* What is wrong with the jailer's reasoning?

Solution. Let AF (BF) be the event that jailer *says* A (B) to be freed. Let AE or BE or CE be the event that A or B or C to be executed. Then, $P(CE) = 1/3$. but, by the *Bayes formula* ,

$$
\begin{aligned}
P(CE|AF) &= \frac{P(AF|CE)P(CE)}{P(AF|AE)P(AE) + P(AF|BE)P(BE) + P(AF|CE)P(CE)}\\
&= \frac{0.5 * 1/3}{0 * 1/3 + 1 * 1/3 + 1/2 * 1/3}\\
&= 1/3\\
&= P(CE).
\end{aligned}
$$

Likewise $P(CE|BF) = P(CE)$. So the "rise of probability" is false.