Lectures 7-8

Example 0.1 *(Ballot problem) In an election between two candidates, candidate I got n votes and II got 10 votes, where $n > 10$. Find the probability that the winner was leading throughout the ballot counting. The above Ballot problem is the simplest of the Ballot problems from Combinatorics.*

For k, l non negative integers, let A_{kl} denote the event that I is leading throughout the counting and F_{l+kk} denote the event that the last ballot was casted for I when I has k votes and II has l votes. Also denote $P(A_{kl}) = p(k,l)$. Observe that

$$p(k,l) = 0, l \geq k, \ p(k,0) = 1 \text{ for all } k, l.$$

Then using Law of total probability, we get

$$P(A_{kl}) = P(A_{kl}|F_{l+kk})P(F_{l+kk}) + P(A_{kl}|F^c_{l+kk})P(F^c_{l+kk}).$$

Now

$$\begin{aligned} P(F_{l+k,k}) &= \frac{\text{no. of ways puting } k-1 \text{ votes in } l+k-1 \text{ locations}}{\text{no. of ways puting } k \text{ votes in } l+k \text{ locations}} \\ &= \frac{\binom{l+k-1}{k-1}}{\binom{l+k}{k}} \\ &= \frac{k}{l+k}. \end{aligned}$$

Hence

$$P(F^c_{l+k,k}) = \frac{l}{l+k}.$$

Now $P(A_{kl}|F_{l+k,k})$ can be viewed as the probability that I is leading from 1 to $l+k-1$ ballot counts when I has $k-1$ votes and II has k votes. Hence

$$P(A_{kl}|F_{l+k,k}) = p(k-1,l).$$

Similarly $P(A_{kl}|F^c_{l+kk})$ can be viewed as the probability that I leading from 1 to $l+k-1$ ballot counts where I has k votes and II has $l-1$ votes. Hence

$$P(A_{kl}|F^c_{l+kk}) = p(k, l-1).$$

Hence we get the following recursive relation

$$P(k,l) = \frac{k}{l+k}p(k-1,l) + \frac{l}{l+k}p(k,l-1), k > l$$

with the boundary conditions $p(k,l) = 0, k \leq l,\ p(k,0) = 1$ for all $k = 0,1,\cdots,n,\ l = 0,1,\cdots,10$. *From the above, we have*

$$\begin{aligned}
p(1,0) &= 1,\ p(1,1) = 0, \\
p(2,0) &= 1,\ p(2,1) = \frac{1}{3},\ p(2,2) = 0, \\
p(3,0) &= 1,\ p(3,1) = \frac{1}{2},\ p(3,2) = \frac{1}{5} \\
p(4,0) &= 1,\ p(4,1) = \frac{3}{5},\ p(4,2) = \frac{1}{3},\ p(4,3) = \frac{1}{7}, p(4,4) = 0.
\end{aligned}$$

Hence, we guess that

$$p(k,l) = \frac{k-l}{k+l}, k \geq l, l+k \geq 2.$$

Now assume that the above is true for $l+k = 0,1,2,\cdots,m$, *then for* (k,l) *with* $l+k = m+1$, *it immediately follows from the recursive relation that*

$$p(k,l) = \frac{k-l}{l+k}, k \geq l, l+k = m+1.$$

This implies by induction that the required probability is $\frac{n-10}{n+10}$.

0.1 Notion of Statistical Independence

Concept of statistical independence (in short independence) is a fundamental concept in probability. We always see statements like "toss a coin repeatedly in an independent fashion' or 'independent trials of a random experiment' etc. What does the above statements mean or how does one give a precise meaning to the word 'independence' appearing in those sentences. To this end, let us look the definition of cartesian product from a different angle. One can observe that given two sets A and B with n and m elements, the cartesian product $A \times B$ is obtained through independent selection, in the sense that choice of $a \in A$ doesnot play any role in selecting $b \in B$. Also note that $\#(A \times B) = \#A \cdot \#B$. This put together implies the following. 'independent selections multiply the number of possibilities'!

The above discussion points that one way of formalizing the notion of independent is using the product rule for probabilities as follows.

Definition 0.1 *(Independence) Two events A, B are said to be independent if*

$$P(AB) = P(A)\,P(B)\,.$$

Remark 0.1 *If $P(A) > 0$, then A and B are independent (i.e. the product rule for probabilities) iff $P(B|A) = P(A)$. This confirms the intuition behind the notion of independence, "the occurrence of one event doesn't have any effect on the probability of the occurrence of the other". i.e., the product rule for probabilities is a good model for the notion of independence.*

Example 0.2 *Define the probability space $(\Omega, \mathcal{F}, P)$ as follows.*

$$\Omega = \{HH, HT, TH, TT\},\ \mathcal{F} = \mathcal{P}(\Omega)$$

and

$$P(\{\omega\}) = \frac{1}{4},\ \omega \in \Omega\,.$$

Consider the events $A = \{HH, HT\}$, $B = \{HH, TH\}$ and $C = \{HT, TT\}$. Then A and B are independent and B and C are dependent.

The notion of independence defined above can be extended to independence of three or more events in the following manner.

Definition 0.2 *(Independence of three events). The events A, B, C are independent (mutually) if*
(i) A, B ; B, C and C, A are independent and
(ii) $P(ABC) = P(A)\,P(B)\ P(C)$.

If the events A, B, C satisfies only (i), then A, B, C are said to be pairwise independent.

Example 0.3 *Define $\Omega = \{1, 2, 3, 4\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and*

$$P(\{i\}) = \frac{1}{4},\ i = 1, 2, 3, 4\,.$$

Consider the events

$$A = \{1, 2\},\ B = \{1, 3\},\ C = \{1, 4\}\,.$$

Then A, B, C are pairwise independent but not independent. One type of dependency.

Example 0.4 *Define* $(\Omega, \mathcal{F}, P)$ *as follows.* $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ *and*

$$P(\{i\}) = \frac{1}{8},\ i = 1, \ldots, 8\,.$$

Let

$$A = \{1,\ 2,\ 3,\ 4\},\ B = \{3,\ 4,\ 5,\ 6\},\ C = \{1, 3,\ 5,\ 7\}\,.$$

Then A, B, C *are independent.*

Example 0.5 *Let* Ω *be given by the 'tossing an unbiased coin thrice' and all sample points are equally likely. Consider the events*

$$\begin{aligned} A &= \{HHH, HHT, HTH, HTT\}, \\ B &= \{HHT, HHH, THT, THH\}, \\ C &= \{HTH, HHH, TTH, TTT\}. \end{aligned}$$

Now see that (exercise)

$$P(ABC) = P(A)P(B)P(C),\ P(BC) \neq P(B)P(C).$$

This is another type of dependency.

Definition 0.3 *The events* $\{A_1, A_2, \ldots, A_n\} \subseteq \mathcal{F}$ *are said to be independent if for any distinct* $A_{i_1},\ \ldots, A_{i_m}, m \geq 2$ *from* $\{A_1,\ A_2, \ldots, A_n\}$

$$P(A_{i_1} \ldots A_{i_m}) = P(A_{i_1}) \ldots P(A_{i_m})\,.$$

We often consider random experiments like 'toss a coin repeatedly countably infinite times in an independent way'. To understand its meaning, we need to extend the definition of independence to denumerable (i.e. countably infinite) number of random variables. To this end we first introduce the notion of independence when there are denumerable number of events.

Definition 0.4 *The events* $\{A_1, A_2, \cdots\} \subseteq \mathcal{F}$ *are said to be independent if any finite subcollection of events from* $\{A_1, A_2, \cdots\}$ *are independent, i.e. for any* $m \geq 1, \{i_1, \imath_2, \cdots, i_m\} \subseteq \mathbb{N}$, $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$ *are independent.*

In the above definition, we are not including the product rule for probabilities for infinite collection, because in practice independent trials are performed sucessively and hence it is only meaning full/ 'practical' to talk about independent among finite collection (note given n independent trials to one can perform a trial which is independent of trials already occured), also it is sufficient in view of continuity property of probabilities.

Example 0.6 *Consider the random experiment of 'tossing an unbiased coin countably infinite times', i.e. any finite collection of tosses are independent. Let A_n denote the event that nth toss results in a H for $n \geq 1$. Then $\{A_1, A_2, \cdots\}$ are independent (exercise-would like to see how you write down the solution!)*

One can define conditional independence using product rule for probabilities with conditional probabilities.

Definition 0.5 *(Conditional independence) We say that two events $A, B \in \mathcal{F}$ are conditionally independent given the event $C \in \mathcal{F}, P(C) > 0$, if*

$$P(AB|C) \;=\; P(A|C)\; P(B|C).$$

Example 0.7 *There are two coins one is a $\frac{1}{6}$-coin[1] and another is a $\frac{5}{6}$-coin. The experiment is the following. Pick a coin at random and toss the selected coin twice independently. Consider the events A, first toss gives a H, B denote the second gives a H and C denote first coin is picked. Now A and B are dependent but A and B are conditionally independent with respect to C (exercise : Check - $P(A \cap B) = \frac{13}{36}, P(A) = P(B) = \frac{1}{2}$)*

Using the notion of independence of events, one can define the independence of σ-fields as follows.

Definition 0.6 *Two σ-fields $\mathcal{F}_1,\ \mathcal{F}_2$ of subsets of Ω are said to be independent if for any $A \in \mathcal{F}_1,\ B \in \mathcal{F}_2$*

$$P(AB) \;=\; P(A)\, P(B)\,.$$

Definition 0.7 *The σ-fields $\mathcal{F}_1,\ \mathcal{F}_2, \ldots \mathcal{F}_n$ of subsets of Ω are said to be independent if $A_1, A_2, \ldots, A_n$ are independent whenever $A_i \in \mathcal{F}_i,\ i = 1, 2, \ldots, n$.*

One can go a step further to define independence of any family of σ-fields as follows.

Definition 0.8 *A family of σ-fields $\{\mathcal{F}_i | i \in I\}$, where I is an index set, are independent if for any finite subset $\{\alpha_1, \cdots \alpha_n\} \subseteq I$, the σ-fields $\mathcal{F}_{\alpha_1}, \cdots, \mathcal{F}_{\alpha_n}$ are independent.*
In particular, when $I = \mathbb{N}$, we get the definition of independence for a countably infinite collection.

[1] We use p-coin to denote a coin with probability p for geting H

Finally one can introduce the notion of independence of random variables through the corresponding σ-field generated. It is natural to define independence of random variables using the corresponding σ-fields, since the σ-field generated by a random variable gives a collective information about the random variable.

Definition 0.9 *Two random variables X and Y are independent if $\sigma(X)$ and $\sigma(Y)$ are independent.*

Definition 0.10 *A family of random variables $\{X_i \mid i \in I\}$ are independent if $\{\sigma(X_i) \mid i \in I\}$ are independent.*

Example 0.8 *Let A, B are independent events iff $\sigma(A)$ and $\sigma(B)$ are independent σ-fields iff the random variables I_A and I_B are independent.*

Proof. *Let A and B are independent. Consider*

$$P(AB^c) \;=\; P(A) \,-\, P(AB) \,=\, P(A)(1 - P(B)) \,=\, P(A)\, P(B^c)\,.$$

Hence A and B^c are independent. Changing the roles of A and B we have A^c and B are independent. Now A^c and B are independent implies that A^c and B^c are independent.
Since

$$\sigma(A) \;=\; \{\emptyset,\, A,\, A^c, \Omega\},\;\; \sigma(B) \;=\; \{\emptyset,\, B,\, B^c, \Omega\}\,,$$

it follows that $\sigma(A)$ and $\sigma(B)$ are independent. Converse statement is obvious.

The second part follows from

$$\sigma(A) \,=\, \sigma(I_A),\;\; \sigma(B) \,=\, \sigma(I_B)\,.$$

Example 0.9 *The trivial σ-field $\mathcal{F}_0 \;=\; \{\emptyset,\, \Omega\}$ is independent of any σ-field of subsets of Ω.*

Example 0.10 *Define a probability space $(\Omega,\, \mathcal{F},\, P)$ as follows:*

$$\Omega \;=\; \{(a_1,\, a_2,\, \ldots,\, a_n) \mid a_i \;=\; 0,\, 1;\;\, i = 1, \ldots, n\},\;\; \mathcal{F} \;=\; \mathcal{P}(\Omega)$$

and

$$P\{(a_1,\, a_2,\, \ldots,\, a_n)\} \;=\; \frac{1}{2^n} \text{ for all } (a_1,\, \ldots,\, a_n) \,\in \Omega\,.$$

For $i = 1, 2, \cdots, n,$ *set*

$$A_i = \{(a_1, a_2, \cdots, a_n) | a_i = 1\}.$$

Then $A_1, A_2, \cdots, A_n$ *are independent.*

This can be seen from the following.

$$P(A_i) = \frac{2^{n-1}}{2^n} = \frac{1}{2} \ \forall \ i = 1, 2, \cdots, n.$$

Note that

$$A_{i_1} \ldots A_{i_m} = \{(a_1, \ldots, a_m) \,|\, a_{i_1} = \cdots = a_{i_m} = 1\}.$$

Hence

$$P(A_{i_1} \ldots A_{i_m}) = \frac{2^{n-m}}{2^n} = \frac{1}{2^m} \ \forall \ i_1 \ldots, i_m.$$

Also

$$P(A_{i_1}) \cdots P(A_{i_m}) = \overbrace{\frac{1}{2} \cdots \frac{1}{2}}^{m \ times} = \frac{1}{2^m}.$$

Example 0.11 *Let* $(\Omega, \mathcal{F}, \ P)$ *is given by as*

$$\Omega = \{HH, HT, TH, TT\}, \ \mathcal{F} = \mathcal{P}(\Omega)$$

and

$$P(\{\omega\}) = \frac{1}{4}, \ \omega \in \Omega.$$

Define two random variables $X_1, \ X_2$ *as*

$$\begin{aligned} X_1(HH) &= X_1(HT) = 1, \ X_1(TH) = X_1(TT) = 0, \\ X_2(HH) &= X_2(TH) = 1, \ X_2(HT) = X_2(TT) = 0. \end{aligned}$$

Then X_1 *and* X_2 *are independent. Here note that*

$$\sigma(X_1) = \{\emptyset, \ \{HH, HT\}, \ \{TH, TT\}, \ \Omega\}$$

and

$$\sigma(X_2) = \{\emptyset, \ \{HH, TH\}, \ \{HT, TT\}, \ \Omega\}.$$

0.1.1 Borel-Cantelli Lemma

In many situations which involves independent trials, one may be interested in knowing the probabilities of the events in which some specific scenario happens repeatedly, for example in a denumerable(countably infinite) independent tosses of a coin, getting infintely many 'HH'. Borel-Cantelli Lemma which will prove next will help us in calculating probabilities of such events. We begin with the set theoretic model for events of these form, which are given by the notion of limsup of sets which is defined below. Borel-Cantelli lemma is then about giving the probability of limsup of events.

Definition 0.11 ($\limsup$ *of sets) For* $A_1, A_2, \cdots,$ *subsets of* Ω *, define*

$$\limsup_n A_n \ := \ \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \, .$$

Similarly

$$\liminf_n A_n \ = \ \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} \ .$$

In the following theorem, we list some useful properties of limsup and liminf. This will make the objects $\liminf$ and $\limsup$ of sets much clearer.

Theorem 0.1 *1. Let*

$$\{A_n \ i.o\} \ := \ \{\omega \in \Omega \,|\, \text{there exists } n_1 < n_2 < \cdots \text{ such that } \omega \in A_{n_k} \ \forall \ k \geq 1\}.$$

Then

$$\limsup_n A_n \ = \ \{A_n \ i.o\} \, .$$

2. Let

$$\{A_n \text{ a.a.}\} \ := \ \{\omega \in \Omega \,|\, \omega \in A_k \text{ for all } k \geq n \text{ for some } n \geq 1\} \, .$$

Here a.a. stands for 'all most all'. Then

$$\liminf_n A_n \ = \ \{A_n \text{ a.a.}\} \, .$$

3. The following inclusion holds.

$$\limsup_n A_n \supseteq \liminf_n A_n \, .$$

4. *The following identity holds.*

$$[\limsup_n A_n]^c \;=\; \liminf_n A_n^c \,.$$

5. *If* $A_1 \subseteq A_2 \subseteq \cdots$, *then*

$$\limsup_n A_n \;=\; \cup_{n=1}^{\infty} A_n \;=\; \liminf_n A_n$$

6. *If* $A_1 \supseteq A_2 \supseteq \cdots$, *then*

$$\limsup_n A_n \;=\; \cap_{n=1}^{\infty} A_n \;=\; \liminf_n A_n \,.$$

Proof.

1. Consider

$$\begin{aligned} \omega \in \limsup_n A_n \quad & \text{iff} \quad \omega \in \cup_{k=n}^{\infty} A_k \text{ for all } n \geq 1 \\ & \text{iff} \quad \text{there exists } n_1 < n_2 < \ldots, \text{ such that } \omega \in A_{n_k}, \text{ for all } k \geq 1 \\ & \text{iff} \quad \omega \in \{A_n \; i.o.\} \end{aligned}$$

This proves (1).

2. Consider

$$\begin{aligned} \omega \in \liminf_n A_n \quad & \text{iff} \quad \omega \in \cap_{k=n}^{\infty} A_k \text{ for some } n \neq 1 \\ & \text{iff} \quad \text{there exists } n_0 \geq 1 \text{ such that } \omega \in A_n, \text{ for all } n \geq n_0 \\ & \text{iff} \quad \omega \in \{A_n \; a.a\} \end{aligned}$$

Thus (2).

3. Proof of (3) follows from (1) and (2).

4. Proof of (4) follows from De Morgen's laws.

5. Note

$$\begin{aligned} \limsup_n A_n \quad \subseteq \quad & \limsup_{n\to\infty} A \,, \text{ where } A = \cup_n A_n \\ & = A. \end{aligned}$$

Hence

$$\limsup_{n\to\infty} A_n \subseteq \cup_n A_n.$$

Now using the fact that A_n's are increasing, we get

$$\liminf A_n \;=\; \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \;=\; \cup_{n=1}^{\infty} A_n .$$

Hence

$$\liminf A_n \;=\; \cup_{n=1}^{\infty} A_n \;\subseteq \limsup_n A_n \subseteq \cup_n A_n A_n .$$

From this (5) follows.

6. The proof of (6) follows from (4) and (5).

Remark 0.2 *The properties (3), (4) and (6) are analogous to the corresponding properties of limsup and liminf of real numbers.*

Remark 0.3 *Analogous to the notion of limit of sequence of numbers, one can say that* $\lim_{n\to\infty} A_n$ *exists if*

$$\liminf_{n\to\infty} A_n \;=\; \limsup_{n\to\infty} A_n \, .$$

Hence from Theorem 0.1 (5), if $\{A_n\}$ *is an increasing sequence of sets, then* $\lim_{n\to\infty} A_n$ *exists and is* $\cup_{n=1}^{\infty} A_n$. *Similarly using Theorem 0.1 (6), if* $\{A_n\}$ *is an decreasing sequence of sets, then* $\lim_{n\to\infty} A_n$ *exists and is* $\cap_{n=1}^{\infty} A_n$. *Now student can see why property (6) in Theorem 1(Lecturenotes 3-4) is called continuity property of probability.*

Theorem 0.2 *(Borel - Cantelli Lemma) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $A_1, A_2, \cdots \in \mathcal{F}$.

(i) If

$$\sum_{n=1}^{\infty} P(A_n) \;<\; \infty$$

then

$$P(\limsup_n A_n) \;=\; 0 \, .$$

(ii) If

$$\sum_{n=1}^{\infty} P(A_n) \;=\; \infty$$

and

$$\{A_n \,|\, n = 1, 2, \cdots\}$$

are independent, then

$$P(\limsup A_n) \;=\; 1 \, .$$

Proof: (i) Consider

$$\begin{aligned} P(\limsup_n A_n) &= P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) \\ &= \lim_{n\to\infty} P(\cup_{k=n}^{\infty} A_k) \\ &\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) \,. \end{aligned}$$

Note that the r.h.s $\to 0$ as $n \to \infty$, since

$$\sum_{n=1}^{\infty} P(A_n) \;<\; \infty \,.$$

Therefore

$$P(\limsup_n A_n) \;=\; 0 \,.$$

(ii) Note that

$$(\limsup_n A_n)^c \;=\; \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c \,.$$

Therefore

$$P(\limsup_n A_n) \;=\; 1 \;-\; P(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) \,.$$

Hence it is enough to show that

$$P(\liminf_{n\to\infty} A_n^c) \;=\; 0.$$

The proof the above uses the following two results.

- $$\{A_1,\, A_2,\, \cdots\} \text{ are independent } \;\Rightarrow\; \{A_1^c,\, A_2^c,\, \cdots\} \text{ are independent} \,.$$

- For each $x \in \mathbb{R}$,
  $$1 - x \leq e^{-x}.$$
  (The proof of this inequality follows from the fact that $f(x) = e^{-x} + x - 1, x \in \mathbb{R}$ has a minimum at $x = 0$ which is unique because f is strictly convex.)

Now

$$\begin{aligned} P(\cap_{k=n}^{\infty} A_k^c) &\leq P(\cap_{k=n}^{n+m} A_k^c) \text{ for all } m = 1,\, 2,\, \cdots \\ &= \Pi_{k=n}^{n+m} P(A_k^c)\,,\ m \geq 1 \end{aligned}$$

The last equality follows from

$$\{A_1,\, A_2,\, \cdots\} \text{ are independent } \Rightarrow \{A_1^c,\, A_2^c,\, \cdots\} \text{ are independent}\,.$$

Using $1 - x \leq e^{-x}$, we get

$$\begin{aligned} P(\cap_{k=n}^{\infty} A_k^c) &\leq \Pi_{k=n}^{n+m}(1 - P(A_k) \\ &\leq \Pi_{k=n}^{n+m} e^{-P(A_k)} \\ &= e^{-\sum_{k=n}^{n+m} P(A_k)},\ m \geq 1. \end{aligned}$$

Hence by letting $m \to \infty$ above , we get

$$P(\cap_{k=n}^{\infty} A_k^c) \leq e^{-\sum_{k=n}^{\infty} P(A_k)},\ n \geq 1.$$

Since

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

we get

$$e^{-\sum_{k=n}^{\infty} P(A_k)} = 0\,,\ n \geq 1.$$

Therefore

$$P(\cap_{k=n}^{\infty} A_k^c) = 0 \ \forall\ n \geq 1.$$

Thus

$$P(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) \leq \sum_{n=1}^{\infty} P(\cap_{k=n}^{\infty} A_k^c) = 0\,.$$

Therefore

$$P(\limsup_n A_n) = 1\,.$$

This completes the proof.