Lectures 5-6

Lemma 0.1 *Let* $\mathcal{I}_1 = \{(-\infty, x] \,|\, x \in \mathbb{R}\}$. *Then* $\sigma(\mathcal{I}_1) = \mathcal{B}_{\mathbb{R}}$

Proof. For $x \in \mathbb{R}$,

$$(-\infty, x] = (x, \infty)^c \in \mathcal{B}_{\mathbb{R}} .$$

Therefore

$$\mathcal{I}_1 \subseteq \mathcal{B}_{\mathbb{R}} .$$

Hence

$$\sigma(\mathcal{I}_1) \subseteq \mathcal{B}_{\mathbb{R}} . \tag{0.1}$$

Now we prove the reverse inclusion. For $b \in \mathbb{R}$,

$$(-\infty, b) = \bigcup_{n=1}^{\infty} \Big(-\infty, \; b - \frac{1}{n} \Big] \in \sigma(\mathcal{I}_1).$$

Also, for $a, b \in \mathbb{R}, \; a < b$,

$$(a, b) = (-\infty, b) \setminus (-\infty, a] \in \sigma(\mathcal{I}_1) .$$

Therefore, $\sigma(\mathcal{I}_1)$ contains all open intervals for the form (a, b). Since any open set in $\mathbb{R}$ can be written as a (countable) union of open intervals with rational end points,[1] it follows that

$$\mathcal{O} \subseteq \sigma(\mathcal{I}_1) ,$$

where $\mathcal{O}$ denote the set of all open sets in $\mathbb{R}$. Thus,

$$\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{O}) \subseteq \sigma(\mathcal{I}_1) . \tag{0.2}$$

Combining (0.1) and (0.2), we get

$$\sigma(\mathcal{I}_1) = \mathcal{B}_{\mathbb{R}}.$$

□

Theorem 0.1 *Let* $X : \Omega \to \mathbb{R}$ *be a function. Then* X *is a random variable defined on* $(\Omega, \mathcal{F}, P)$ *iff* $X^{-1}(B) \in \mathcal{F}$ *for all* $B \in \mathcal{B}_{\mathbb{R}}$.

[1] Recall that from the definition of open set, given any element of the open set there is an open interval containg the element contained in the open set. Now using the property that ' between any two reals there is a rational' we can choose another open interval with rational end points containing the element and contained in the open set.

Proof. Suppose X is a random variable. i.e.,

$$\{X \leq x\} \;=\; X^{-1}\Big((-\infty,\, x]\Big) \in \mathcal{F} \text{ for all } x \in \mathbb{R}. \tag{0.3}$$

Set

$$\mathcal{D} \;=\; \{B \in \mathcal{B}_{\mathbb{R}} \,|\, X^{-1}(B) \in \mathcal{F}\}\,.$$

From (0.3), we have

$$\mathcal{I}_1 \subseteq \mathcal{D} \subseteq \mathcal{B}_{\mathbb{R}}. \tag{0.4}$$

Note that $X^{-1}(\mathbb{R}) \;=\; \Omega$. Hence $\mathbb{R} \in \mathcal{D}$.
Now

$$\begin{aligned} B \in \mathcal{D} \;&\Rightarrow\; X^{-1}(B) \in \mathcal{F} \\ &\Rightarrow\; [X^{-1}(B)]^c \in \mathcal{F} \\ &\Rightarrow\; X^{-1}(B^c) \in \mathcal{F} \text{ (since } [X^{-1}(B)]^c \;=\; X^{-1}(B^c)] \\ &\Rightarrow\; B^c \in \mathcal{D}\,. \end{aligned}$$

Also

$$\begin{aligned} B_1,\, B_2,\, \cdots \in \mathcal{D} \;&\Rightarrow\; X^{-1}(B_n) \in \mathcal{F}\ \forall\ n \;=\; 1,\, 2,\, 3,\, \cdots \\ &\Rightarrow\; \cup_{n=1}^{\infty} X^{-1}(B_n) \in \mathcal{F} \\ &\Rightarrow\; X^{-1}(\cup_{n=1}^{\infty} B_n) \in \mathcal{F} \\ &\Rightarrow\; \cup_{n=1}^{\infty} B_n \in \mathcal{D}\,. \end{aligned}$$

Hence $\mathcal{D}$ is a σ-field. Now from (0.4) and Lemma 0.1, it follows that $\mathcal{D} \;=\; \mathcal{B}_{\mathbb{R}}$. i.e, $X^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{B}_{\mathbb{R}}$.

Converse statement follows from the observation

$$\mathcal{I}_1 \subseteq \mathcal{B}_{\mathbb{R}}\,.$$

This completes the proof.

Remark 0.1 *Theorem 0.1 tells that if* $X^{-1}\Big((-\infty,\; x]\Big) \in \mathcal{F}$ *for all* $x \in \mathbb{R}$*, then* $X^{-1}(B) \in \mathcal{F}$ *for all* $B \in \mathcal{B}_{\mathbb{R}}$*.*

Example 0.1 *Let* $\Omega \;=\; (0,\, 1]$, $\mathcal{F} \;=\; \mathcal{B}_{(0,\,1]}$*. Define* $X(\omega) \;=\; 3\omega \;+\; 1$*. Here* $\mathcal{B}_{(0,\,1]}$ *is the* σ*-field generated by all open sets in* $(0, 1]$[2]

For $x \in \mathbb{R}$

$$\begin{aligned} \{X \leq x\} \;&=\; \{\omega \in \Omega \,|\, 3\,\omega \;+\; 1 \leq x\}\,. \\ &=\; \begin{cases} \emptyset & \text{if} \quad x \;<\; 1 \\ (0,\, \frac{1}{3}x \;-\; \frac{1}{3}] & \text{if} \quad 1 \leq x \leq 4 \\ (0,\;1] & \text{if} \quad x \;>\; 4\,. \end{cases} \end{aligned}$$

[2]A set O in $(0,\; 1]$ is open if $O = (0,\; 1] \cap G$, for some G which is an open set in $\mathbb{R}$.

Since

$$\emptyset,\ (0,\ \frac{1}{3}x\ -\ \frac{1}{3}],\ (0,\,1] \in \mathcal{B}_{(0,\,1]}\,,$$

X is a random variable.

Lemma 0.2 *For a non empty set Ω and a map $X : \Omega \to \mathbb{R}$, define*

$$\sigma(X)\ =\ \{X^{-1}(B)\,|\,B \in \mathcal{B}_{\mathbb{R}}\}\,.$$

Then $\sigma(X)$ is a σ-field.

Proof.

$$X^{-1}(\mathbb{R})\ =\ \Omega\,.$$

Hence $\Omega \in \sigma(X)$.

For $A \in \sigma(X)$, there exists $B \in \mathcal{B}_{\mathbb{R}}$ such that

$$X^{-1}(B)\ =\ A$$

and

$$A^c\ =\ X^{-1}(B^c)\,.$$

Hence $A \in \sigma(X)$ implies $A^c \in \sigma(X)$. Similarly from

$$X^{-1}(\cup_{n=1}^{\infty} B_n)\ =\ \cup_{n=1}^{\infty} X^{-1}(B_n)$$

it follows that

$$A_1, A_2, \cdots \in \sigma(X) \Rightarrow \cup_{n=1}^{\infty} A_n \in \sigma(X)\,.$$

This completes the proof.

Definition 0.1 *For a non empty set Ω and a map $X : \Omega \to \mathbb{R}$, $\sigma(X)$ is called the* **σ-field generated by** *X.*

Remark 0.2 *(1)The occurrence or nonoccurence of the event $X^{-1}(B)$ tells respectively whether any realization $X(\omega)$ is in B or not. Thus $\sigma(X)$ collects all such information. Hence $\sigma(X)$ can be viewed as the information available with the random variable.*

(2) For a function X defined on a sample space Ω, $\sigma(X)$ is a σ-field of subsets of Ω and is the smallest σ-field under which X is a random variable. A map $X : \Omega \to \mathbb{R}$ is a random variable with respect to a σ-field $\mathcal{F}$ iff $\sigma(X) \subseteq \mathcal{F}$.

Example 0.2 *Let X be as in Example 5 (Lecturenotes3-4), i.e.* $\Omega = \{1,2,3,4,5,6\}$ *and* $X = I_{\{1\}}$*. Then for* $B \in \mathcal{B}_{\mathbb{R}}$*,*

$$X^{-1}(B) \;=\; \begin{cases} \emptyset & \text{if} \quad B \cap \{0,1\} = \emptyset \\ \{2,3,4,5,6\} & \text{if} \quad B \cap \{0,1\} = \{0\} \\ \{1\} & \text{if} \quad B \cap \{0,1\} = \{1\} \\ \Omega & \text{if} \quad B \cap \{0,1\} = \{0,1\} \end{cases}$$

Hence $\sigma(X) = \sigma(\{1\})$*. It is easy to see that X is a random variable with respect to any σ-field containing* $\sigma(\{1\})$ *and* $\sigma(\{\{2\},\{4\},\{6\},\{1,3,5\}\})$ *doesn't contain* $\sigma(\{1\})$*.*

Given two random variables X, Y on Ω with respect to a σ-field $\mathcal{F}$, and $c \in \mathbb{R}$, we define sum, scalar product and product of X, Y, i.e. $X + Y : \Omega \to \mathbb{R}, cX : \Omega \to \mathbb{R}, XY : \Omega \to \mathbb{R}$ as follows.

$$(X+Y)(\omega) = X(\omega)+Y(\omega), (cX)(\omega) = cX(\omega),\ (XY)(\omega) = X(\omega)Y(\omega),\ \omega \in \Omega.$$

Theorem 0.2 *If X, Y are random variables on $(\Omega, \mathcal{F})$ and $c \in \mathbb{R}$, then $X + c$, cX, $X + Y$, X^2 and XY are random variables with respect to $\mathcal{F}$.*

Proof: The proof of first two are simple exercises.
For $a \in \mathbb{R}$,

$$\{X + Y < a\} \;=\; \bigcup_{r \in \mathbb{Q}} \big(\{X < r\} \cap \{Y < a - r\}\big) \in \mathcal{F}\,.$$

The proof of the above identity, only uses "between any two real numbers, there is always a rational" [3] Now

$$\{X + Y \le a\} \;=\; \bigcap_{n=1}^{\infty} \Big\{X + Y < a + \frac{1}{n}\Big\}\,.$$

Hence $\{X + Y \le a\} \in \mathcal{F}$ for all $a \in \mathbb{R}$. Therefore $X + Y$ is a random variable.

For $a \in \mathbb{R}$,

$$\{X^2 \le a\} \;=\; \begin{cases} \emptyset \in \mathcal{F} & \text{if} \quad a < 0 \\ \{X = 0\} = X^{-1}(\{0\}) \in \mathcal{F} & \text{if} \quad a = 0 \\ \{-\sqrt{a} \le X \le \sqrt{a}\} = X^{-1}([-\sqrt{a},\ \sqrt{a}]) \in \mathcal{F} & \text{if} \quad a > 0\,. \end{cases}$$

[3]Hint: For $X(\omega) + Y(\omega) < a$, choose r, a rational number such that $X(\omega) < r < a - Y(\omega)$. Then $\omega \in \{X < r\} \cap \{Y < a - r\}$.

(In the above $\{0\}$ and $[-\sqrt{a},\ \sqrt{a}]$ are closed sets and hence Borel sets.) Hence, X^2 is a random variable.

Note that

$$XY \;=\; \frac{1}{2}[(X+Y)^2 \;-\; X^2 \;-\; Y^2]\,.$$

Since $\frac{1}{2}(X\,+\,Y)^2$, $-\frac{1}{2}X^2$, $-\frac{1}{2}Y^2$ are random variables and XY is their sum, XY is a random variable.

Theorem 0.3 *(i) If X, Y are random variables with respect to $\mathcal{F}$, so are* $\min\{X,\,Y\}$, $\max\{X,\,Y\}$*.*

(ii) If $\{X_n\}$ is a sequence of random variables on $(\Omega, \mathcal{F})$ which is bounded from above and

$$X(\omega) \;=\; \sup_n X_n(\omega)\ ,\ \text{for all } \omega \in \Omega.$$

Then X is a random variable with respect to $\mathcal{F}$. Analogous statement is true for infimum.

Proof:

(i) Set $Z \;=\; \min\{X,\,Y\}$. For $a \in \mathbb{R}$,

$$\{Z \le a\} \;=\; \{X \le a\} \cup \{Y \le a\} \in \mathcal{F}\,.$$

Therefore Z is a random variable.
The proof of $\max\{X,\,Y\}$ is a random variable is similar.

Proof of (ii) follows from

$$\{X \le a\} \;=\; \cap_{n=1}^{\infty}\{X_n \le a\}\,.$$

Theorem 0.4 *Let $\{X_n\}$ be a sequence of random variables on $(\Omega,\,\mathcal{F})$ such that*

$$\lim_{n\to\infty} X_n(\omega)\ \text{ exists for all } \omega \in \Omega\,.$$

Then $X : \Omega \to \mathbb{R}$ defined by

$$X(\omega) \;=\; \lim_{n\to\infty} X_n(\omega)\,,\ \omega \in \Omega$$

is a random variable.

Proof. For $a \in \mathbb{R}$.

$$\begin{aligned} \omega \in \{X \leq a\} \quad &\Rightarrow \quad \text{for each } m \geq 1, \text{there exists } n \text{ such that} \\ &\qquad X_k(\omega) \leq X(\omega) \ + \ \tfrac{1}{m} \ \leq \ a + \tfrac{1}{m} \text{ for all } k \geq n \\ &\Rightarrow \quad \omega \in \textstyle\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_k \leq a + \tfrac{1}{m} \right\}, m \geq 1 \\ &\Rightarrow \quad \omega \in \textstyle\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \cap_{k=n}^{\infty} \left\{ X_k \leq a + \tfrac{1}{m} \right\}. \end{aligned}$$

Hence

$$\{X \leq a\} \subseteq \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_k \leq a + \frac{1}{m} \right\}.$$

Now suppose

$$\omega \in \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_k \leq a + \frac{1}{m} \right\}.$$

If $\omega \notin \{X \leq a\}$, then there exists m_0 and n_0 such that

$$X_k(\omega) > a \ + \ \frac{1}{m_0} \text{for all } k \geq n_0 \,. \tag{0.5}$$

(This follows, since $X_n(\omega) \to X(\omega) > a$ as $n \to \infty$)
Also

$$\begin{aligned} \omega \in \textstyle\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_k \leq a + \tfrac{1}{m} \right\} \quad &\Rightarrow \quad \omega \in \textstyle\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_k \leq a \ + \tfrac{1}{m_0} \right\} \\ &\Rightarrow \quad X_k(\omega) \leq a \ + \ \tfrac{1}{m_0} \\ &\qquad \text{for all } k \geq n_1 \text{ for some } n_1 \end{aligned}$$

This contradicts (0.5). Therefore

$$\omega \in \{X \leq a\} \,.$$

Hence we have

$$\{X \leq a\} \ = \ \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_k \leq a \ + \ \frac{1}{m} \right\}. \tag{0.6}$$

Since $\{X_k \in a + \frac{1}{m}\} \in \mathcal{F}$ and $\mathcal{F}$ is a σ-field, using(0.6) it follows from the definition of σ-field that

$$\{X \leq a\} \in \mathcal{F} \,.$$

Therefore X is a random variable.

Definition 0.2 *A function $f : \mathbb{R} \to \mathbb{R}$ is said to be Borel (measurable) function if $f^{-1}(B) \in \mathcal{B}_{\mathbb{R}}$ for all $B \in \mathcal{B}_{\mathbb{R}}$.*

Example 0.3 *If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then it is Borel function. This follows from the fact that $f^{-1}(O)$ is open if O is open in $\mathbb{R}$.*

Example 0.4 *If $f : \mathbb{R} \to \mathbb{R}$ is a monotone function (i.e. either increasing[4] or decreasing), then f is a Borel function.*

Let f be increasing. For $a \in \mathbb{R}$, consider the set $D_a := \{f > a\}$. Note that in general D_a need not be bounded from below, see for example $f(x) = e^x, x \in \mathbb{R}$ and $\{f > 0\} = \mathbb{R}$ itself.

Set

$$c = \inf\{x | f(x) > a\}.$$

There are two cases: $c \in D_a$ or $c \notin D_a$.

When $c \in D_a$ (then $c > -\infty$).

$$x \geq c \Rightarrow f(x) \geq f(c) > a \Rightarrow x \in D_a.$$

Hence $[c, \infty) \subseteq D_a$. From the definition of D_a and c, it follows that $D_a \subseteq [c, \infty)$. Hence $D_a = [c, \infty) \in \mathcal{B}_{\mathbb{R}}$.

When $c \notin D_a$ (then it is possible that $c = -\infty$).

$$x > c \Rightarrow \text{ there exists } y \in D_a \text{ such that } c < y < x \Rightarrow f(x) \geq f(y) > a \Rightarrow x \in D_a.$$

Hence $(c, \infty) \subseteq D_a$. Also it follows that $D_a \subseteq (c, \infty)$. Therefore $D_a = (c, \infty) \in \mathcal{B}_{\mathbb{R}}$. Thus we have

$$D_a = \begin{cases} \mathbb{R} & \text{if } D_a \text{ is not bounded from below} \\ [c, \infty) & \text{if } c = \inf D_a \in D_a \\ (c, \infty) & \text{if } c \notin D_a. \end{cases}$$

i.e. $D_a = \{f > a\}$ is an interval. Hence $\{f > a\} \in \mathcal{B}_{\mathbb{R}}$ for all $a \in \mathbb{R}$. From this we can see that $\{f \leq a\} \in \mathcal{B}_{\mathbb{R}}$ for all $a \in \mathbb{R}$. Therefore f is a Borel function.

Now we have a result which gives us plenty of random variables.

[4] f is increasing if when ever $x < y$, then $f(x) \leq f(y)$

Lemma 0.3 *Let $f : \mathbb{R} \to \mathbb{R}$ be a Borel function and $X : \Omega \to \mathbb{R}$ is a random variable with respect to $\mathcal{F}$. Then $f \circ X$ is a random variable with respect to $\mathcal{F}$.*

Proof. For $B \in \mathcal{B}_{\mathbb{R}}$,

$$(f \circ X)^{-1}(B) \;=\; X^{-1}(f^{-1}(B)) \in \mathcal{B}_{\mathbb{R}},$$

since $f^{-1}(B) \in \mathcal{B}_{\mathbb{R}}$. Hence $f \circ X$ is a random variable.

Example 0.5 *The natural logarithm*[5] $\log : (0,\ \infty) \to \mathbb{R}$ *is defined by*

$$\log x \;=\; \int_1^x \frac{1}{t} dt.$$

Note that $\log$ *is a strictly increasing and continuous function which is also a bijective map. We define the exponential function $e : \mathbb{R} \to (0,\ \infty)$ as the inverse of* $\ln$. *Then both are Borel measurable and hence* $\log X,\ e^X$ *are random variables if X is a random variable.*

[5]Natural logarithm of positive number is invented by John Napier in 1614. The definition we are using is the logarithmic propety of the quadrature (determining area under a curve) rule for hyperbola. The notation log was introduced by Leibnitz in 1675 and subsequently in 1676, he invented the logarithmic property of the quadrature of hyperbola.

Chapter 3: Conditional Probability and Independence

In this chapter, we introduce the concepts of conditional probability and independence.

"Conditioning" is an important technique for computing probabilities. Conditioning is about writing down the probability of an event in terms of probabilities of the event given that some other events have already occurred. For using the above method, one need to understand these 'new' probabilities. Let us take an example to illustrate this. Suppose person ONE is inside a room and throw a die and secretly tell his friend TWO that an even number turned up. Now the person THREE ask TWO how probable was the event $\{1,2,3\}$? TWO will rule out the outcomes 1 and 3 and scale it using even numbers to arrive at the new probability $\frac{1}{3}$ which is different from $\frac{1}{2}$. i.e., information changes the probabilities, these are called conditional probabilities. Formalizing the above 'slicing' and 'scaling' leads to the following definition of conditional probability.

Definition 4.1 Let $(\Omega, \mathcal{F}, P)$ be a probability space and $A \in \mathcal{F}$ be such that $P(A) > 0$. The conditional probability of an event $B \in \mathcal{F}$ given A denoted by $P(B|A)$ is defined as

$$P(B|A) \;=\; \frac{P(AB)}{P(A)}\,. \tag{0.7}$$

Define P_A on $\mathcal{F}$ as follows.

$$P_A(B) \;=\; P(B|A),\; B \in \mathcal{F}\,.$$

Then $(\Omega,\ \mathcal{F}, P_A)$ is a probability space satisfying

$$P_A(A) = 1,\; P_A(B) = 0 \text{ if } B \subseteq A^c.$$

(exercise)

Remark 0.3 *The above probability space may look useless but it is useful for processing the thought experiments of the following type. Note that $(\Omega, \mathcal{F}, P_A)$ corresponds to a random experiment which is obtained by adding the information that the event A has occurred. Hence by understanding directly this random experiment gives the probabilities $P_A(B)$ directly without using the formula (0.7) for $P_A(B)$. We will see an illustration of this through examples.*

Example 0.6 *Let* $\Omega = \{1,2,3,4,5,6\}$ *and* $\mathcal{F} = \mathcal{P}(\Omega)$ *and* $P(\{i\}) = \frac{1}{6}$ *for all* i. *Suppose, we got the information that 'even number' has occured, then what are the new 'probabilities' i.e., the conditional probabilities of the events?*

Set $A = \{2,4,6\}$. *We need to compute* P_A. *Then from (0.7), we have*

$$P_A(\{i\}) = \frac{1}{3},\ i = 2,4,6, \text{ and } = 0,\ i = 1,3,5.$$

Example 0.7 *(Bridge hand) A pack of* 52 *cards are distributed among four players and is called a bridge hand. Find the probability of a balanced bridge hand of aces, i.e. each player gets an ace.*

Define the following events. $A_{\clubsuit}$ *denote the event that a player gets the ace of club,* $A_{\clubsuit,\diamondsuit}$, *the event that two distinct players get an ace of club and diamond each,* $A_{\clubsuit,\diamondsuit,\heartsuit}$, *the event that three distinct players get an ace of club, diamond and heart and* $A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit}$, *the event that all players get an ace. Then*

$$A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit} \subset A_{\clubsuit,\diamondsuit,\heartsuit} \subset A_{\clubsuit,\diamondsuit} \subset A_{\clubsuit}.$$

and hence using $A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit} = A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit} \cap A_{\clubsuit,\diamondsuit,\heartsuit} \cap A_{\clubsuit,\diamondsuit} \cap A_{\clubsuit}$, *we get (exercise)*

$$P(A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit}) = P(A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit}|A_{\clubsuit,\diamondsuit,\heartsuit})P(A_{\clubsuit,\diamondsuit,\heartsuit}|A_{\clubsuit,\diamondsuit})P(A_{\clubsuit,\diamondsuit}|A_{\clubsuit})P(A_{\clubsuit}).$$

(The above is a conditioning argument)
Now $P(A_{\clubsuit}) = 1$. *To compute*[6] $P(A_{\clubsuit,\diamondsuit}|A_{\clubsuit})$, *observe that given* $A_{\clubsuit}$, *distributing diamond ace to one of the remaining three players is equivalent to placing the diamond ace in one of the 39 locations out of a total of 51 locations. Hence*

$$\begin{aligned} P(A_{\clubsuit,\diamondsuit}|A_{\clubsuit}) &= \frac{\text{no. of ways diamond ace can be placed in 39 locations}}{\text{no. of ways diamond ace can be placed in 51 locations}} \\ &= \frac{\binom{39}{1}}{\binom{51}{1}} = \frac{39}{51}. \end{aligned}$$

[6]To compute the conditional probabilities, we will not use the definition of conditional probability instead we use an equivalent random experiment which gives raise to the probability space of conditional probabilities, as told in Remark 0.3. Note that given the information that the club ace is distributed to one player, we can think about the random experiment as an urn problem with 51 urns numbered 1 to 51 and a ball (identified with diamond ace) and our event is distributing the ball into one of the first 39 urns.

Similarly

$$P(A_{\clubsuit,\diamondsuit,\heartsuit}|A_{\clubsuit,\diamondsuit}) \;=\; \frac{26}{50},\;\; P(A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit}|A_{\clubsuit,\diamondsuit,\heartsuit}) \;=\; \frac{13}{49}.$$

Hence

$$P(A_{\clubsuit,\diamondsuit,\heartsuit,\spadesuit}) \;=\; \frac{39 \times 26 \times 13}{51 \times 50 \times 49} \;=\; 0.11.$$

The above example illustrates that it is some times more easy to compute (or natural to specify) conditional probabilities and use them to specify the underlying probabilities, a reverse procedure!

Example 0.8 *(Probability in the game show "Let's make a deal") Here we look at a version of the game show "Let's make a deal" which made its debut on NBC Television network on December 30, 1963. Description of the game is the following. A prize is placed behind one of the three doors and are closed. Contestants of the show are aware of this. Contestant is asked to select a door (but is not going to open at the moment). Once the choice is made, the moderator of the show (Monty Hall) opens one of the remaining doors and display what is in it (He will only open a door which has no prize). Contestant now is given a chance to change the earlier choice. Question is, will the contestant stay with the earlier choice or not?*

Without any loss of generality, assume that the contestant chose door no.1 (label chosen door as no.1).

Here take sample space as

$$\Omega = \{\diamondsuit ab, \diamondsuit ba, a\diamondsuit b, b\diamondsuit a, ab\diamondsuit, ba\diamondsuit\}.$$

(Here $\diamondsuit ab$ denotes price $\diamondsuit$ behind the first door and the 'worthless' a and b behind the doors 2 and 3 respectively. Other sample points are similraly interpreted.)

Question can be answered if we know the probability of '$\diamondsuit$ behind door 1' given the additional information of the object behind one of the doors 2 or 3.

Let us denote the event '$\diamondsuit$ behind door 1' by '$\diamondsuit \in 1$'. Also Monty Hall revealing object behind door 2 means behind door 2, object is either a or b. Hence occurrence of the event A, i.e. revealing door 2 means occurrence of $\{\diamondsuit ab, \diamondsuit ba, ab\diamondsuit, ba\diamondsuit\}$. i.e.

$$A \;=\; \{\diamondsuit ab, \diamondsuit ba, ab\diamondsuit, ba\diamondsuit\}.$$

Similarly the event of 'revealing door 3' is given the occurrence of

$$B \; := \; \{\Diamond ab, \Diamond ba, a\Diamond b, b\Diamond a\}.$$

Hence we need to find $P(\Diamond \in 1 | A \cup B)$. *Since* $A \cup B = \Omega$, *we get*

$$P(\Diamond \in 1 | A \cup B) \;=\; P(\Diamond \in 1) = \frac{1}{3}.$$

Hence it is better to change the option, since you have 2/3rd chance of winning the prize by switching the door.

We have already seen 'finite' partitions of the sample space. A collection $\{A_1, A_2, \ldots, A_N\}$ of events is said to be a countable *partition of* Ω, if

(i) A_i's are pairwise disjoint

(ii) $\cup_{i=1}^N A_i \;=\; \Omega\,.$

Here N may be ∞. If $N < \infty$, then partition is said to be finite partition and if $N = \infty$[7], it is called a countably infinite partition.

Theorem 0.5 *(Law of total probability-discrete form) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $\{A_1, A_2, \ldots, A_N\} \subseteq \mathcal{F}$ *be a countable partition of* Ω *such that* $P(A_i) > 0$ *for all* i. *Then for* $B \in \mathcal{F}$,

$$P(B) \;=\; \sum_{i=1}^{N} P(B|A_i)\, P(A_i)\,.$$

Proof. Consider

$$\begin{aligned}
\sum_{i=1}^{N} P(B|A_i)\, P(A_i) &= \sum_{i=1}^{N} \frac{P(BA_i)}{P(A_i)} P(A_i) \\
&= \sum_{i=1}^{N} P(BA_i) \\
&= P(B(\cup_{i=1}^N A_i)) \;=\; P(B)\,.
\end{aligned}$$

The second last equality uses the countable additivity of probability when $N = \infty$ to get convergence of the series.

[7]When $N = \infty$, the collection is $\{A_1, A_2, \cdots\}$

Theorem 0.6 *(Bayes Theorem) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $A,\, B \in \mathcal{F}$ *such that* $P(A),\, P(B) > 0$ *and* $P(A) < 1$. *Then*

$$P(A|B) \;=\; \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}\,.$$

Proof.

$$P(A|B) \;=\; \frac{P(BA)P(A)}{P(A)\,P(B)} \;=\; \frac{P(B|A)P(A)}{P(B)}\,.$$

Now use Law of total probability to complete the proof.

Remark 0.4 *The total probability law is another formula for using conditioning argument.*

Example 0.9 *(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.*