Course Notes

Introduction to Probability

Alex Rutar

BSM Fall 2018

Contents

Chapter 1

Fundamentals

1.1 Basic Principles

1.1.1 Probability Spaces

A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$.

1.1.2 Ω

Ω is a set, called the sample space, and $\omega \in \Omega$ are called outcomes and $A \subset \Omega$ are called events.

Ex. 1.1.1 A horserace with 3 horses, a, b, c, has $\Omega = \{(a,b,c),(a,c,b),\ldots,(c,b,a)\}$. Then $|\Omega| = 6$ and $A = \{a \text{ wins the race}\} = \{(a,b,c),(a,c,b)\}$.

Ex. 1.1.2 Roll two fair dice, a white die and a yellow die. Then $\Omega = \{(1,1),(1,2),\ldots,(6,6)\}$ and $|\Omega| = 36$.

Ex. 1.1.3 Continue flipping a coin until there is a head. Then

$$\Omega = \{(H),(T,H),(T,T,H),\ldots\}$$

Then define

$$A = \{\text{there are an even number of rolls}\} = \{(T,H),(T,T,T,H),\ldots\}$$

Ex. 1.1.4 Consider $\Omega = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 100\}$. Then $A = \{\text{you score 50 points}\} = \{(x,y) \mid x^2+y^2 \leq 1\}$.

Def'n. 1.1.5 *If $A \cap B = \emptyset$, we say that A and B are* ***mutually exclusive*** *events. If $A \subset B$, we say that A* ***implies*** *B.*

Write $A^c = \Omega \setminus A$. Recall distributivity, the deMorgan relations, etc.

1.1.3 $\mathcal{F}$

$\mathcal{F}$ is a collection of subsets of Ω, which denote the events that we consider.

- If Ω is countable, then typically $\mathcal{F}$ is just the collection of all subsets of Ω.
- If Ω is a domain in $\mathbb{R}^n$, then it is a strict subset of $\mathbb{R}^n$.

In any case, $\mathcal{F}$ has to be closed under the following operations:

1. $\Omega \in \mathcal{F}$
2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
3. If $A_1, A_2, \ldots \in \mathcal{F}$, then $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

in other words, that $\mathcal{F}$ is a $\sigma-$algebra.

1.1.4 $\mathbb{P}$

Finally, $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is a function that satisfies 3 axioms:

1. For any $A \in \mathcal{F}$, then $\mathbb{P}(A) \geq 0$
2. $\mathbb{P}(\Omega) = 1$
3. ($\sigma-$additivity) Let $A_1, A_2, A_3, \ldots$ be a sequence of mutually exclusive events. Then

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

1.1.5 Consequences

- $\mathbb{P}(A^c) + \mathbb{P}(A) = \mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1$.
- If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$ since $\mathbb{P}(B) = \mathbb{P}((A^c \cap B) \cup (A \cap B)) = \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B) = \mathbb{P}(A^c \cap B) + \mathbb{P}(A)$
- For any A, B, we have

  $$\mathbb{P}(A \cup B) = \mathbb{P}((A^c \cap B) \cup (A \cap B) \cup (A \cap B^c)) = \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B) + \mathbb{P}(B^c \cap A) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

  Similarly,

  $$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

  which generlizes arbitrarily:

  $$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{1 \leq i_1 < i_2 < \cdots < i_r \leq n} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_r})$$

Proof We have already proved the base case for $n = 2$, so assume the formula holds for a union of n events. Then

$$\mathbb{P}(A_1 \cup \cdots A_n \cup A_{n+1}) = \mathbb{P}(A_1 \cup \cdots \cup A_n) + \mathbb{P}(A_{n+1}) - \mathbb{P}((A_1 \cup \cdots \cup A_n) \cap A_{n+1})$$

We can distribute the first and third terms using the induction hypothesis, and the result follows. □

Def'n. 1.1.6 *We say $D_1, D_2, \ldots$ is a* ***decreasing*** *sequence of events of $D_{k+1} \subset D_k$. We say $D_1, D_2, \ldots$ is a* ***increasing*** *sequence of events of $D_{k+1} \supset D_k$.*

Let $\lim_{n\to\infty} D_n = \bigcap_{n=1}^\infty {}_n$ and $\lim_{n\to\infty} I_n = \bigcup_{n=1}^\infty I_n$.

Prop. 1.1.7 *$\sigma-$additivity implies that for any increasing sequence,*

$$\Pr\left(\lim_{n\to\infty} I_n\right) = \lim_{n\to\infty} \Pr(I_n)$$

and similarly for any decreasing sequence

$$\Pr\left(\lim_{n\to\infty} D_n\right) = \lim_{n\to\infty} \Pr(D_n)$$

Proof Note that (2) implies (1): if D_k is a decreasing sequence, then $I_k = D_k^c$ is an increasing sequence and

$$\left(\lim_{n\to\infty} D_n\right)^c = \left(\bigcap_{n=1}^\infty D_n\right)^c = \bigcup_{n=1}^\infty I_n = \lim_{n\to\infty} I_n$$

and taking probabilities,

$$\Pr\left(\lim_{n\to\infty} D_n\right) = 1 - \Pr\left(\lim_{n\to\infty} I_n\right) = 1 - \lim_{n\to\infty} \Pr(I_n) = \lim_{n\to\infty} \Pr(D_n)$$

To prove that $\sigma-$additivity implies (1), let $I_1, I_2, \ldots$ be increasing. Let $A_1 = I_1$ and for $k \geq 2$ let $A_k = I_k \setminus I_{k-1}$. Then $A_1, A_2, \ldots$ are mutually exclusive and for any $k \geq 1$,

$$\bigcup_{k=1}^K A_k = I_k$$

Thus

$$\bigcup_{k=1}^\infty A_k = \lim_{n\to\infty} I_n$$

Now note that $\Pr(I_K) = \sum_{k=1}^K \Pr(A_k)$ while

$$\begin{aligned}\Pr\left(\lim_{n\to\infty} I_n\right) &= \Pr\left(\bigcup_{k=1}^\infty A_k\right) \\ &= \sum_{k=1}^\infty \Pr(A_k) \\ &= \lim_{K\to\infty} \sum_{k=1}^K \Pr(A_k) \\ &= \lim_{K\to\infty} \Pr(I_K)\end{aligned}$$ □

1.1.6 Examples with Finite Uniform Probabilities

We assume that $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$ and $\mathbb{P}(\{\omega_i\}) = \mathbb{P}(\{\omega_j\})$. Then $\mathbb{P}(\{\omega_i\}) = \frac{1}{N}$ and $\mathbb{P}(A) = |A|/N$.

Ex. 1.1.8 In an urn there are 6 blue balls and 5 red balls. Draw 3 balls out of this 11. What is the change that among the 3 there are exactly 2 blue balls and 1 red ball?

Let us pretend that the balls are labelled, 1 through 11, and set Ω to be all the ordered triples of disjoint elements. Then $A = \{\text{exactly 2 blue and 1 red}\}$, and note that $A = A^1 \cup A^2 \cup A^3$ where A^i has a red in position i and blue in the other two positions. Now, $|A^i| = 5 \cdot 6 \cdot 5$, so $|A| = 3 \cdot 6 \cdot 5 \cdot 6$ and

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3 \cdot 6 \cdot 5 \cdot 6}{11 \cdot 10 \cdot 9}$$

We now suppose that $\Omega = \{\Lambda \subset \{1, \dots, 11\} \mid |\Lambda| = 3\}$, so $|\Omega| = \binom{11}{3}$. Now

$$A = \{\Lambda_1 \cup \Lambda_1 | \Lambda_1 \subset \{1, \dots, 6\}, |\Lambda_1| = 2, \Lambda_2 \subset \{7, \dots, 11\}, |\Lambda_2| = 1\}$$

So $|A| = \binom{6}{2} \cdot 5$.

Ex. 1.1.9 Consider a group of N people. What is the chance that there is at least one pair amoung them who have the same birthday?

Define $\Omega = \{(i_1, i_2, \dots, i_N) \mid i_j \in \{1, \dots, 365\}\}$. We want $A = \{\text{there is at least one common birthday}\}$. We can write

$$A^c = \{(i_1, \dots, i_n) \in \Omega \mid i_j \neq i_k \forall j \neq k\}$$

Then $|A^c| = 365 \cdot 364 \cdots (365 - N + 1)$ and

$$P_N = \mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{365 \cdot 364 \cdots (365 - N + 1)}{365^N}$$

Ex. 1.1.10 Suppose we have N people at a party. The following day, everyone leaves one after another, and chooses a single phone from a pile. What is the chance that nobody chooses her own phone?

Define $\Omega = \{(i_1, \dots, i_N) \mid \text{permutations of } \{1, \dots, N\}\}$, so $\omega = (i_1, \dots, i_k)$ means person k chooses phone i_k. Then $|\Omega| = N!$. Fix $B = \{\text{nobody picks her/his phone}\}$. Define $A_1 = \{\text{person 1 picks his phone}\}$, so $|A_1| = (N-1)!$, and similarly for A_2, etc. Then $B = A_1^c \cap A_2^c \dots \cap A_N^c = (A_1 \cup \dots \cup A_N)^c$, and $\mathbb{P}(A_i) = \frac{1}{N}$. Now in general,

$$\Pr(A_{i_1} \cap \dots \cap A_{i_k}) = \frac{(N-k)!}{N!}$$

for i_k distinct. Thus we now have

$$\begin{aligned}
\Pr(B) &= 1 - \Pr(A_1 \cup A_2 \cup \dots \cup A_n) \\
&= 1 - \sum_{r=1}^{N} (-1)^{r+1} \sum_{1 \leq i_1 < i_2 \dots < i_r \leq N} \Pr(A_{i_1} \cap \dots \cap A_{i_r}) \\
&= \sum_{r=1}^{n} (-1)^{r+1} \binom{N}{r} \frac{(N-r)!}{N!} \\
&= \sum_{r=1}^{N} (-1)^{r+1} \frac{1}{r!}
\end{aligned}$$

so that
$$\Pr(B) = 1 + \sum_{r=1}^{N} (-1)^r \frac{1}{r!} = \sum_{r=0}^{N} (-1)^r \frac{1}{r!}$$
Thus $\lim_{N\to\infty} \Pr(B) = \frac{1}{e}$.

Ex. 1.1.11 (Round table seating) Consider a round table with 20 seats, and 10 married couples sit. What is the change that no couples sit together?

Define $\Omega = \{\text{permutations of } \{1,\ldots,20\}/\sim\}$ where $(i_1,\ldots,i_{20}) \sum (i_{20}, i_1, \ldots, i_{19})$. Then $|\Omega| = 19!$. Define $B = \{\text{no couples together} = A_1^c \cap A_2^c \cap \cdots \cap A_{10}^c\}$, where
$$A_k = \{\text{the 8th woman sits next to her spouse}\}$$
so that
$$\Pr(B) = 1 - \Pr(A_1 \cup \cdots \cup A_{10})$$
Note that
$$\Pr(A_i) = \frac{18!2}{19!} = \frac{2}{19}$$
by "joining" the couple together, arranging them around the table, and permuting the couple internally. Thus generalizes to
$$\Pr(A_{i_1} \cap \cdots \cap \Pr(a_{i_r}) = \frac{2^r(19-r)!}{19!}$$
Then by inclusion-exclusion,
$$\Pr(B) = 1 - \binom{10}{1} \cdot \frac{18!2}{19!} + \binom{10}{2}\frac{17!2^2}{19!} - \binom{10}{3}\frac{16!2^3}{19!} \cdots + \binom{10}{10}\frac{9!2^{10}}{19!} \approx 0.339$$

Ex. 1.1.12 (Poker hand probabilities) A poker hand is a straight if the 5 cards are of increasing value and not all of the same suit, starting with $A, 2, 3, 4, \ldots, 10$.

Define $\Omega = \{\text{5 element subsets of the 52 cards}\}$. Then $|\Omega| = \binom{52}{5}$. Thus
$$\Pr(\text{straight}) = \frac{10 \cdot (4^5 - 4)}{\binom{52}{5}}$$
$$\Pr(\text{full house}) = \frac{13 \cdot 12 \cdot \binom{4}{3} \cdot \binom{4}{2}}{\binom{52}{5}}$$

Ex. 1.1.13 (Bridge hand probabilities) In bridge, each of the 4 players get 13 cards. Let $\Omega = \{\text{13 cards that North gets}\}$.
$$\Pr(\text{North receives all spaces}) = \frac{1}{\binom{52}{13}}$$
Pr(North does not receive all 4 suits of any value) = Pr(There is some value such that all suits are at N)

Let $V_k = \{\text{North gets all four suits of value } k\}$. Then

$$\Pr(V_1) = \frac{\binom{48}{9}}{\binom{52}{13}}$$

$$\Pr(V_1 \cap V_2) = \frac{\binom{44}{5}}{\binom{52}{13}}$$

$$\Pr(V_1 \cap V_2 \cap V_4) = \frac{\binom{40}{1}}{\binom{52}{13}}$$

Thus

$$1 - \Pr(V_1 \cup V_2 \cup \cdots \cup V_{13}) = 1 - \frac{\binom{48}{9}}{\binom{52}{13}} \cdot 13 + \binom{13}{2}\frac{\binom{44}{5}}{\binom{52}{13}} - \binom{13}{3}\frac{40}{\binom{52}{5}}$$

What is the change that each player receives one ace? There are

$$\frac{52!}{13!13!13!13!}$$

possible hands. There are 4! ways to arrange the aces, which gives

$$\Pr(E) = \frac{4!\binom{48}{12,12,12,12}}{\binom{52}{13,13,13,13}}$$

1.2 Conditional Probability

1.2.1 Basic Principles

Suppose we roll two fair dice. Then Pr(the sum is 10) $= \frac{3}{36} = \frac{1}{12}$. Suppose instead that the white dice is rolled first, and it turns up 6. Now the probability that the sum is 10 is now 1/6.

Def'n. 1.2.1 *Given an even E with $\Pr(E) > 0$, for any event F, let $\Pr(F|E) = \frac{\Pr(F\cap E)}{\Pr(E)}$. We call this the **conditional probability of** F **given** E.*

Prop. 1.2.2 *Fix E with $\Pr(E) > 0$ and consider $\Pr(\cdot|E) : \mathcal{F} \to \mathbb{R}$. This function satisfies the axioms of probability.*

Proof
1. $\Pr(F|E) \geq 0$ for all $F \in \mathcal{F}$.

2. $\Pr(\Omega|E) = \frac{\Pr(E\cap\Omega)}{\Pr(E)} = 1$

3. If $F_1, F_2, \ldots$ are mutually exclusive, then

$$\begin{aligned}
\Pr(\bigcup_{i=1}^{\infty} F_i | E) &= \frac{\Pr((\bigcup_{i=1}^{\infty} F_i) \cap E)}{\Pr(E)} \\
&= \frac{\Pr(\bigcup_{i=1}^{\infty} (E \cap F_i))}{\Pr(E)} \\
&= \sum_{n=1}^{\infty} \frac{\Pr(F_i \cap E)}{\Pr(e)} \\
&= \sum_{n=1}^{\infty} \Pr(F_n | E)
\end{aligned}$$ □

Prop. 1.2.3 *We have* $\Pr(E \cap F) = \Pr(F|E) \cdot \Pr(E)$, *and more generally*

$$\Pr(E_n \cap E_{n-1} \cap \cdots \cap E_1) = \Pr(E_n | E_{n-1} \cap \cdots \cap E_1) \cdots \Pr(E_3 | E_2 \cap E_1) \Pr(E_2 | E_1) \Pr(E_1)$$

Proof This follows by induction from the definition of conditional probability. □

Ex. 1.2.4 Andrew and Bob play for the college basketball team. They get two T-shirts each, in closed bags. Any T-shirt can be black or white, with 50-50 chance. Andrew prefers black, but Bob has no preference. The following day, Andrew shows up with a black shirt on. What is the chance that Andrew's other shirt is black?

Sol'n We have $\Omega = \{(B,B),(B,W),(W,B),(W,W)\}$ which is reduced to $\{(B,B),(B,W),(W,B)\}$, so the answer is 1/3. To make this transparent, consider

$$\begin{aligned}
A_1 &= \{\text{Andrew has at least one black shirt}\} \\
A_2 &= \{\text{Both of Andrew's shirts are black}\} \\
A_3 &= \{\text{Andrew has a black shirt on}\}
\end{aligned}$$

so in Andrew's case, $A_1 = A_3$ and $\Pr(A_2|A_3) = \Pr(A_2|A_1)$.

Ex. 1.2.5 (Polya's Urn) Initially, we have two balls, 1 red, 1 blue, in the urn. For the first draw, pick one, check its color, and put it back and put another ball of the same color into the urn.

1. What is Pr(the first three balls are red, blue, red (in this order)).

Sol'n 1. Let R_i, B_i denote the i^{th} draw is red or blue respectively. Then

$$\Pr(R_3 \cap B_2 \cap R_1) = \Pr(R_3 | B_2 \cap R_1) \Pr(B_2 | R_1) \Pr(R_1) = \frac{1}{2}\frac{1}{3}\frac{1}{2} = \frac{1}{2}$$

Ex. 1.2.6 What is Pr(in bridge, each of the players gets one ace)?

Sol'n Write

$$\begin{aligned}
&E_4\\
&\ \cap\\
&E_3 = \{\text{Aces of spaces, heards, and diamonds are at 3 different players.}\}\\
&\ \cap\\
&E_2 = \{\text{Aces of spaces, hearts, and diamonds are at 2 diferent players.}\}\\
&\ \cap\\
&E_1 = \Omega
\end{aligned}$$

so that $\Pr(E_4) = \Pr(E_4 \cap E_3 \cap E_2 \cap E_1) = \Pr(E_4|E_3)\Pr(E_3|E_2)\Pr(E_2|E_1)\Pr(E_1)$.

1.2.2 Bayes' Formula

Ex. 1.2.7 Consider an insurance compacy, which classifies people into accident prone drivers (30%) and non-accident-prone drivers, (70%). For accident prone drivers, the chance of being involved in an accident within a year is 0.2, while for non-addicent-prone drivers, the chance of being involved in an accident is 0.1. Now suppose we have a new policyholder.

1. What is the probability that the policyholder is involved in an accident within a year?
2. The policyholder was involved in an accident?

Sol'n

1. $B = \{\text{accident in 2018}\}$, $A = \{\text{the policyholder is accident prone}\}$. Then

$$\Pr(B) = \Pr(B\cap A)+\Pr(B\cap A^c) = \Pr(B|A)\Pr(A)+\Pr(B|A^c)\Pr(A^c) = 0.2\cdot 0.3+0.1\cdot 0.7 = 0.13$$

2. Now

$$\Pr(A|B) = \frac{\Pr(A\cap B)}{\Pr(B)} = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A)+\Pr(B|A^c)\cdot\Pr(A^c)} = \frac{0.2\cdot 0.3}{0.13} = \frac{6}{13}$$

Prop. 1.2.8 *Suppose $A_1, A_2, \dots, A_n \in \mathcal{F}$ form a partition of Ω. Given such a partition, for any $B \in \mathcal{F}$,*

$$\Pr(B) = \sum_{i=1}^{n} \Pr(B\cap A_i) = \sum_{i=1}^{n} \Pr(B|A_i)\cdot\Pr(A_i)$$

Then for any $k \in [n]$,

$$\Pr(A_k|B) = \frac{\Pr(B\cap A_k)}{\Pr(B)} = \frac{\Pr(B|A_k)\cdot\Pr(A_k)}{\sum\limits_{i=1}^{n}\Pr(B|A_i)\cdot\Pr(A_i)}$$

Ex. 1.2.9 Roll a fair dice. There is a urn with one white ball in it. If the die turns up 1,3, or 5, put one black ball ito the urn. If it turns up 2 or 4, put 3 black and 5 white, and if it turns up 6, put 5 black and 5 white.

Sol'n Write

$$\begin{aligned}
A_1 &= \{1,3 \text{ or } 5 \text{ rolled}\}\\
A_2 &= \{2 \text{ or } 4 \text{ rolled}\}\\
A_3 &= \{6 \text{ rolled}\}B && = \{\text{black ball rolled}\}
\end{aligned}$$

so that

$$\begin{aligned}
\Pr(A_3|B) &= \frac{\Pr(B|A_3)\Pr(A_3)}{\Pr(B|A_1)\cdot\Pr(A_1)+\Pr(B|A_2)\cdot\Pr(A_2)+\Pr(B|A_3)\cdot\Pr(A_3)} \\
&= \frac{5/6\cdot 1/6}{1/2\cdot 1/2+3/4\cdot 1/3+5/6\cdot 1/6} \\
&= \frac{5}{23}
\end{aligned}$$

Ex. 1.2.10 There is a blood test for a rare but serious disease. Only 1/10000 people have this disease. Suppose the test is 100% effective, so if someone is tested ill, it is positive with 100% chance. Suppose there is also a 1% chance of false positive.

A new patient is tested, and tests positive. What are the odds that she has the disease?

Sol'n Let $A = \{\text{the person is ill}\}$ and $B = \{\text{the test is positive}\}$. Then

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A)+\Pr(B|A^c)\Pr(A^c)} = \frac{1\cdot 0.0001}{1\cdot 0.0001+0.01\cdot 0.9999}$$

Ex. 1.2.11 (Monty Hall paradox) There are three doors: one of them hides a prize, and two hide nothing. Pick a door. The announcer then reveals another door not containing a prize. Is it better to stay or switch?

Sol'n Write $A_i =$ ¶door i hides the price, and $B_2 = \{\text{door 2 is opened}\}$. Then

$$\Pr(A_1|B_2) = \Pr(B_2|A_1)\Pr(A_1) = \frac{\Pr(B_2|A_1)\Pr(A_1)}{\Pr(B_2|A_1)\Pr(A_1)+\Pr(B_2|A_2)\Pr(A_2)+\Pr(B_2|A_3)\Pr(A_3)} = \frac{1/2\cdot 1/3}{1/2\cdot 1/3+0+1/}$$

but

$$\Pr(A_3|B_2) = \Pr(B_2|A_3)\Pr(A_3) = \frac{\Pr(B_2|A_1)\Pr(A_1)}{\Pr(B_2|A_1)\Pr(A_1)+\Pr(B_2|A_2)\Pr(A_2)+\Pr(B_2|A_3)\Pr(A_3)} = \frac{1\cdot 1/3}{1/2\cdot 1/3+0+1/}$$

so it is better to switch!