Fourier Analysis and Probability Theory

Adithya Ganesh

October 14, 2019

1 Crash course on foundation of probability

A probability space is a triple $(\Omega, \mathcal{F}, p)$, where

- Ω is an arbitrary set. Every $\omega \in \Omega$ is called an event.
- $\mathcal{F}$ is a σ-algebra, i.e., a collection of subsets of Ω such that: $\Omega \in \mathcal{F}$, if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, if $\{A_i\}_{i=1}^{+\infty}$ are all in $\mathcal{F}$, then $\cup_{i=1}^{+\infty} A_i \in \mathcal{F}$.
- $P : \mathcal{F} \to [0,1]$ is a probability measure, i.e., $P(\Omega) = 1$ and $P(\cup_i A_i) = \sum_i P(A_i)$ for every collection of mutually disjoint subsets A_i.

Remark. *Why σ-algebras are important?*

- *For uncountable Ω's is often impossible to construct a probability measure associated to the largest possible σ-algebra (namely, the power set of Ω) if we want some reasonable properties to be satisfied (for example, if $\Omega = \mathbb{R}$, we'd like $P(a, b) = b - a$.).*
- *We will consider different σ-algebras on the same space: we should think of σ-algebras as "information": the largest the σ-algebra, the largest the information.*

A random variable is a measurable function X from a probability space $(\Omega, \mathcal{F}, P)$ to some measurable space $(S, \mathcal{G})$. Being measurable means that the pre-image of $\mathcal{G}$ via X should lie in $\mathcal{F}$. Classical example if $S = \mathbb{R}$, $\mathcal{G} =${ smallest σ-algebra containing all intervals}.
The σ-algebra generated by X, denoted by $\sigma(X)$, is the collection of all sets of the form $X^{-1}(\mathcal{G})$: it represents the "information carried by X".

Given a random variable X and another random variable Y, what is the best guess for the outcome of X given that I observed Y? Or more generally, given a σ-algebra $\mathcal{F}$, what is the best guess for the next value of X?

Definition. *Given an integrable random variable X (i.e., $E(|X|) < +\infty$, and a σ-algebra $\mathcal{F}$, the conditional distribution of X given $\mathcal{F}$, denoted by $E(X|\mathcal{F})$, is the unique random variable which is* *measurable with respect to $\mathcal{F}$ satisfying*

$$E(X1_A) = E(E(X|\mathcal{F})1_A)$$

for all $A \in \mathcal{F}$.

If $\mathcal{F} = \sigma(Y)$ for some random variable Y, then $E(X|Y)$ is a function of Y (this is what "measurable with respect to $\mathcal{F}$ mean in this case), and it corresponds to my best guess of X given the result of Y.

1.1 Exercises

Prove the following:

- if X and Y are independent, then $E(X|Y) = E(X)$: in other words, the random variable $E(X|Y)$ is deterministic, which makes sense since the best guess I can do after observing Y is the same of the best guess I would do without observing (because of independence).
- If $X = f(Y)$, then $E(X|Y) = X = f(Y)$: my best guess if I observe Y is exactly $f(Y)$. This is the opposite case of before, namely, complete correlation.

1.2 An example

Let $\Omega = \{0,1\}^n, \mathcal{F} = \mathcal{P}(\Omega), P$, where P is the uniform on all outcomes (i.e., $P(A) = \frac{|A|}{2^n}$). Consider the functions $X_i : \Omega \to \mathbb{R}$ given by $X_i(\omega) = \omega_i$, where $\omega = (\omega_i)_{i=1}^n \in \Omega$. In other words, X_i is a random variable which gives "the outcome of coin i". $\mathcal{F}_i := \sigma(X_i)$ is the collection of four subsets: the empty set, the whole set Ω, the subset $A = \{0,1\}^{i-1} \times \{0\} \times \{0,1\}^{n-i}$ (which is the event "coin i came out tail") and the subset A^c (i.e., the event "coin i came out head").

Let $S_j = \sum_{i=1}^j X_i$ be the number of heads up to time j. Let's check using the definition the obvious fact that $E(S_j|S_{j-1}) = S_{j-1} + \frac{1}{2}$. First of all, the function $S_{j-1} + \frac{1}{2}$ is a function of S_{j-1}, so it is measurable with respect to the σ-algebra generated by S_{j-1}. Now, every set $A \in \sigma(S_{j-1})$ is of the form $A = S_{j-1}^{-1}(B)$ for some B subset of the real (actually, of the integers!). Concretely, A could be the event "S_{j-1} is an even number". By linearity of expectation, we can split all such events into the event $\{S_{j-1}^{-1} = k\}$ for various k (this is just saying that the event "S_{j-1} is even" can be viewed as the countable union of $S_{j-1} = 0, 2, 4, ...$

Therefore, we need to check whether it is true that

$$E\Big(S_j 1_{S_{j-1}=k}\Big) = E\Big(\Big(S_{j-1} + \frac{1}{2}\Big)1_{S_{j-1}=k}\Big).$$

The right hand side is equal to $(k + \frac{1}{2})P(S_{j-1} = k)$. The left hand side is equal to

$$\begin{aligned} E\Big(\Big(S_{j-1} + X_j\Big)1_{S_{j-1}=k}\Big) &= kP(S_{j-1} = k) + E(X_j 1_{S_{j-1}=k}) = \\ &= (k + E(X_j))P(S_{j-1} = k) = \\ &= (k + \frac{1}{2})P(S_{j-1} = k) \end{aligned}$$

where I used the exercises above.

2 Martingales and stopping times

Definition. *We say that a sequence $S_0, S_1, S_2, \ldots$ is a martingale if*

$$E(S_n|\mathcal{F}_{n-1}) = S_{n-1},$$

where $\mathcal{F}_n$ is the σ-algebra generated by $S_1, \ldots, S_n$.

In other words, a martingale is a model of a fair game: the best guess given the past is the current situation. By definition of conditional expectation where we take $1_A \equiv 1$ (i.e., $A = \Omega$), one obtains

$$E(S_n) = E(E(S_n|\mathcal{F}_{n-1})) = E(S_{n-1}) = \ldots = E(S_0).$$

Usually by convention S_0 is deterministic (often 0, if it represents the net gain at time n in some gambling problem).

In other words, the expected fortune you have at time n is just the one you have at time 0. Usually in gambling problems, we are interested in the following question: is there a strategy I can use, where I only use information up to the present, that allows me stop and gain money out of this? Notice that I'm not asking $E(S_n) > 0$ (which we know it is impossible by our previous observation), but we are rather saying $E(S_T) > 0$ where T is itself random, but somehow T should only "look at the past". Here is a more formal way to put this:

Definition. *Given a sequence of nested $\mathcal{F}_0, \mathcal{F}_1, \ldots$ of σ-algebras (i.e., $\mathcal{F}_{i-1} \subset \mathcal{F}_i$, this is often referred to as a filtration) a stopping time adapted to the filtration $\mathcal{F}_i$ is a random variable T whose range is in $\mathbb{N} \cup \{\infty\}$ with the property that $\{T = n\}$ is $\mathcal{F}_i$ measurable.*

Informally, this is saying that "deciding whether $T = n$ should only use the information up to time n". The stopping σ-algebra $\mathcal{F}_T$ consists of sets A such that $A \subset \{T = n\} \subset \mathcal{F}_n$.

Doob optional stopping theorem is a result which says that, under some assumption on the stopping rule T, we still have $E(S_T) = E(S_0)$ for martingales, i.e., no betting strategy can allow you to win extra money. One version is the following (there are variation though, hypothesis are not minimal here):

Theorem. *Let $S_0, S_1, S_2, \ldots$ be a martingale. Let T be a stopping time such that $\sup_n |S_{\min(T,n)}| < +\infty$. Then $E(S_T) = E(S_0)$.*

2.1 Example

Here is an example: consider S_n be a sum of n independent random variables X_i such that X_i is equal to 1 or -1 with equal probability. Let $T_{a,b}$ be the first time that the martingale hits either a or $-b$. You should think of a being the target and b your initial capital. What is the probability the you will win, i.e., that $S_T = a$? We can use Doob optional stopping theorem: notice that $\sup_n |S_{\min(T,n)}| \leq \max(a, b)$, so that the hypothesis work. Therefore,

$$E(S_T) = E(S_0) = 0.$$

On the other hand, $E(S_T) = aP_{win} - bP_{loss}$, where P_{win} is the probability of winning and P_{loss} is the probability of losing. Since their sum is one, rearranging we obtain

$$P_{win} = \frac{b}{a+b}.$$

Notice that the hypothesis that $\sup_n |S_{\min(T,n)}| < +\infty$ is not redundant. Consider for example T to be the stopping rule "first time I hit a, I stop". In this case, $|S_{\min T,n}|$ can be arbitrarily large (I may be losing a ton of money). Let's try to forget about that and apply Doob optional stopping theorem. If it were true, $E(S_T) = E(S_0) = 0$. Notice that $S_T = a$ almost surely (somehow technical to prove, the idea is that eventually you will always hit a, you are just not sure of how much you have to wait). Therefore, $E(S_T) = E(a) = a$, which is a contradiction!

2.2 Exercises

Try to do the following:

- Under the same definitions for S_i, define $M_n = S_n^2 - n$. Prove that it M_n is a martingale, and use optional stopping theorem with respect to $T_{a,b}$ to show what is $E(T)$.
- If $X_i = 1$ with probability p and equal to -1 otherwise and $S_n = \sum_{i=1}^n X_i$, show that $M_n := \left(\frac{p}{q}\right)^{S_n}$ is a martingale (convention $M_0 = 1$ this time). Use it to prove what is the probability of win/loss using optional stopping theorem.