Tune-Up: Strong Law of Large Numbers

When we do an experiment over and over – something we don't know and can't quite place a formula to and can't quite figure out. So we build an apparatus and observe the experiment with time. Two common guesses:

- IID
- random variable

In order to identify a random variable, we are making a number of observation and trying to make a description. Then IID abstraction, these random variables were not independent or identicaly distributed. The observations and different times were certainly related.

Thm Let $X_1, X_2, \dots$ be IID random variables with $\mathbb{E}(|X_k|) < \infty$, for all k. Define the "total":

$$S_n := X_1 + X_2 + dots + X_n$$

Then with $\mu := \mathbb{E}(X_k)$ be the expected value (the same for all numbers), define the "average":

$$n^{-1} S_n \to \mu(\text{ almost surely })$$

Here "almost surely" means ...

Let Y_n be defined as in the previous lemma (**Komogorov Truncation Lemma**. By property (ii) of that lemma, it's enough to show

$$n^{-1} \sum_{k \leq n} Y_k \to \mu \text{ (almost surely)}$$

Here Y_n is the truncated random variable (to be less than a certain value).

$$Y_n = \begin{cases} X_n & \text{if } |X_n| \leq n \\ 0 & \text{if } |X_n| > n \end{cases}$$

Example, if X_n is an indicator random variable $X_n = 0$ or 1, then $X_n = Y_n$ always.

$$\mathbb{P}[Y_n = X_n \text{ eventually }] = 1$$

What could "**eventually**" mean here ?

We have that

$$n^{-1}\sum_{k\leq n} Y_k = n^{-1}\sum_{n\leq n}\mathbb{E}(Y_k) + n^{-1}\sum_{k\leq n}\big(Y_k - \mathbb{E}(Y_k)\big)$$

So we look at the truncated random variable, differences from our expectation there and then let $n\to\infty$. We now move to the previous version of the Strong Law.

Lemma Let W_n be a sequence of independnet random variables such that

- $\mathbb{E}(W_n) = 0$ (centered)
- $\sum \frac{\mathsf{Var}(W_n)}{n^2} < \infty$ (instead of $\mathbb{E}(|X_k| < \infty)$).

Then $n^{-1}\sum_{k\leq n} W_k \to 0$ (almost surely).

We needed **Dominated Convergence Theorem** to say that $\mathbb{E}(Y_n)\to\mathbb{E}(X)=\mu$.

Most of the work seems to have been in the Kolmogorov Truncation Lemma. These are great placeholders before a more genuine careful look at sequences-and-series, **martingales** and independent of incrememts, e.g. Pythagoras formula:

$$\mathbb{E}(M_n^2) = \mathbb{E}(M_0^2) + \sum_{k=1}^{n}\mathbb{E}[(M_k - M_{k-1})^2]$$

and L^2-bounded martingales, it's written $\mathcal{L}^2(\Omega,\mathcal{F},\mathbb{P})$ so our choice of observables depends on the choice of probablity measure and also what events we choose to consider "measurable" – what were the starting points for our deductive reasoning?

References

[1] ...