Discussion 11B Recap

Tyler Zhu

April 17, 2020

1 Weak Law of Large Numbers

The idea of the weak law of large numbers is to look at the behavior of say coin flips in the long run. There's two questions we can start off our discussion with: how many heads (mean) will we get, and how variable are our results (variance)?

We can describe it formally as such. We perform an experiment n times independently and note

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (X_i's \text{ i.i.d. })$$

where X_i has mean μ and variance σ^2. Then

$$\mathbb{E}[M_n] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\mathbb{E}[X_1] \cdot n = \mathbb{E}[X_1] = \mu,$$

and

$$\text{Var}(M_n) = \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}(X_i) = \frac{\sigma^2}{n}$$

since the X_i are i.i.d. If we take $n \to \infty$ to look at the long term behavior, $\mathbb{E}[M_n] = \mu$ and $\text{Var}(M_n) = 0$.

This is cool and all, but wouldn't it be dope if we also knew the *rate* at which the variance decreased to 0? We can do this by using Chebyshev to do a tail bound:

$$\mathbb{P}(|M_n - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2},$$

and just like that we've derived the weak law of large numbers.

Theorem 1 (Weak LLN)**.** Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. RVs with mean μ. Then for any $\epsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| < \epsilon\right) \to 1$$

as $n \to \infty$.

What does the Weak LLN really mean? In short, it means that $\lim_{n\to\infty} \mathbb{P}(|M_n - \mu| < \epsilon) = 1$. The probability that the average of our trials deviates from the expectation by more than ϵ is 0 in the long run. This does **not** say that M_n will eventually equal μ; in fact, there are sequences of r.v.'s (X_i) which satisfy the statement of the Weak LLN but do not eventually equal the expectation. We would need to invoke the Strong LLN for that, but that's outside the scope of this course.

This is a plot I made illustrating rolling a lot of 6-sided die. The Weak LLN places a bounding box of width 2ϵ around our mean, and says that the probability of any realizations outside this

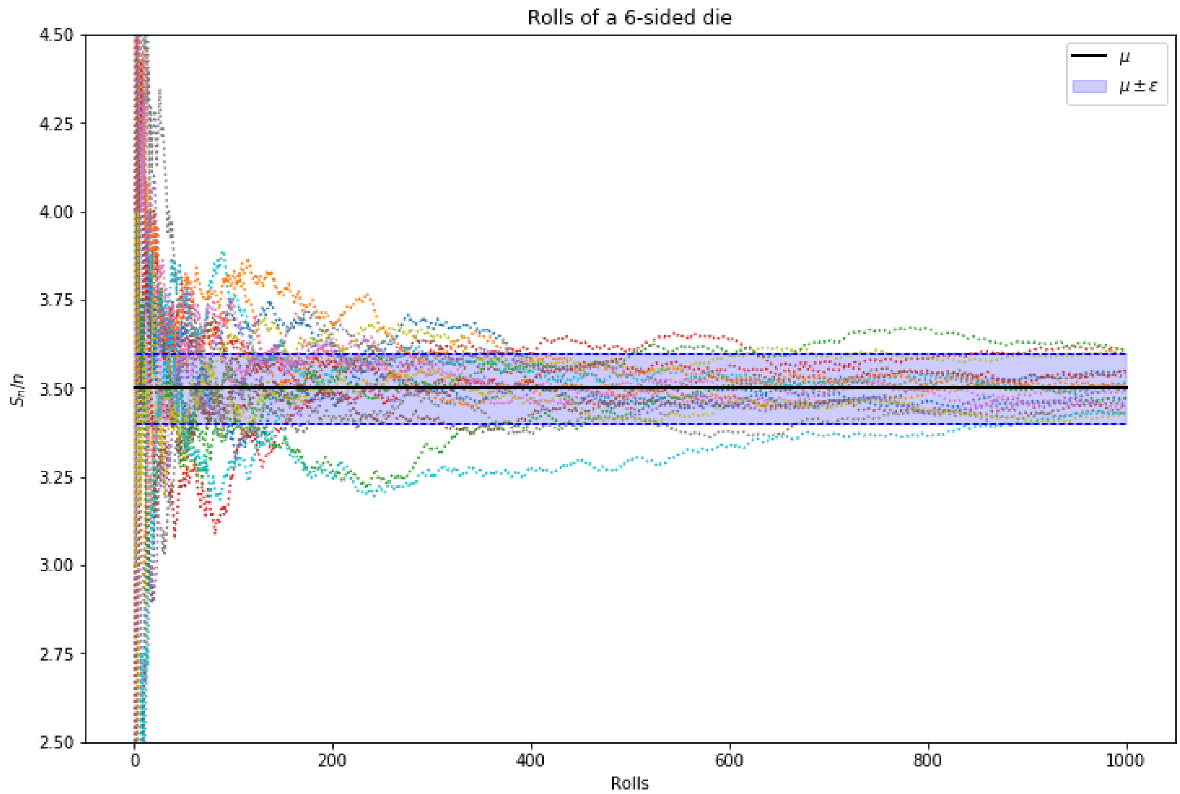

Figure 1: The weak LLN bounds the long-run values with the blue bounding box $X \in [\mu - \epsilon, \mu + \epsilon]$, while strong LLN asserts all values become the black line, $X = \mu$.

box is 0 in the long run. It doesn't say anything about occurences within the box though. Hence, all we are guaranteed is that the *fraction* of realizations outside $\mu \pm \epsilon$ for all $\epsilon > 0$ converges to 0.

2 Confidence Intervals

From the above derivation, we can use Chebyshev's inequality in a similar manner to construct 95% (for example) confidence intervals.