Independent Sampling Theorem

Weak Law of Large Numbers

\[ A_n \triangleq \text{avg of } n \text{ indep RV's with mean } \mu \]

Theorem: For all \( \delta > 0 \)

\[
\lim_{n \to \infty} \Pr[ | A_n - \mu | > \delta ] = 0
\]

Proof:

Repeated Trials

\[
E[A_n] \equiv E \left[ \frac{R_1 + R_2 + \ldots + R_n}{n} \right] = \frac{E[R_1] + E[R_2] + \ldots + E[R_n]}{n} = \frac{n \mu}{n} = \mu
\]
Weak Law of Large Numbers

So by Chebyshev

$$\text{Pr}[| A_n - \mu | > \delta ] \leq \frac{\text{Var}[A_n]}{\delta^2}$$

need only show

$$\text{Var}[A_n] \rightarrow 0 \text{ as } n \rightarrow \infty$$

Analysis of the Proof

proof only used that $R_1, \ldots, R_n$ have

• same mean
• same variance
• & variances add

— which follows from pairwise independence

Var[$A_n$] 

$$\text{Var}[A_n] = \text{Var} \left[ \frac{R_1 + R_2 + \cdots + R_n}{n} \right]$$

$$= \frac{\text{Var}[R_1] + \text{Var}[R_2] + \cdots + \text{Var}[R_n]}{n^2}$$

$$\text{QED} = \frac{n \sigma^2}{n^2} = \frac{\sigma^2}{n} \rightarrow 0$$

Pairwise Independent Sampling

Theorem:

Let $R_1, \ldots, R_n$ be pairwise independent random random vars with the same finite mean $\mu$ and variance $\sigma^2$. Let

$$A_n := \frac{(R_1 + R_2 + \cdots + R_n)}{n}.$$ 

Then

$$\text{Pr}[| A_n - \mu | > \delta ] \leq \frac{1}{n} \left( \frac{\sigma}{\delta} \right)^2$$
Pairwise Independent Sampling

The punchline:
we now know how big a sample is needed to estimate the mean of any* random variable within any* desired tolerance with any* desired probability

*variance < ∞, tolerance > 0, probability < 1