So Chebyshev’s inequality tells us that the final sum will lie between

\[ 7n - 10\sqrt{\frac{35}{6}n} \quad \text{and} \quad 7n + 10\sqrt{\frac{35}{6}n} \]

in at least 99% of all experiments when \( n \) fair dice are rolled. For example, the odds are better than 99 to 1 that the total value of a million rolls will be between 6.976 million and 7.024 million.

In general, let \( X \) be any random variable over a probability space \( \Omega \), having finite mean \( \mu \) and finite standard deviation \( \sigma \). Then we can consider the probability space \( \Omega^n \) whose elementary events are \( n \)-tuples \( (\omega_1, \omega_2, \ldots, \omega_n) \) with each \( \omega_k \in \Omega \), and whose probabilities are

\[ \Pr(\omega_1, \omega_2, \ldots, \omega_n) = \Pr(\omega_1)\Pr(\omega_2)\cdots\Pr(\omega_n) \, . \]

If we now define random variables \( X_k \) by the formula

\[ X_k(\omega_1, \omega_2, \ldots, \omega_n) = X(\omega_k) \, , \]

the quantity

\[ X_1 + X_2 + \cdots + X_n \]

is a sum of \( n \) independent random variables, which corresponds to taking \( n \) independent “samples” of \( X \) on \( \Omega \) and adding them together. The mean of \( X_1 + X_2 + \cdots + X_n \) is \( n\mu \), and the standard deviation is \( \sqrt{n}\sigma \); hence the average of the \( n \) samples,

\[ \frac{1}{n}(X_1 + X_2 + \cdots + X_n) \, , \]

will lie between \( \mu - 10\sigma/\sqrt{n} \) and \( \mu + 10\sigma/\sqrt{n} \) at least 99% of the time. In other words, if we choose a large enough value of \( n \), the average of \( n \) independent samples will almost always be very near the expected value \( EX \). (An even stronger theorem called the Strong Law of Large Numbers is proved in textbooks of probability theory; but the simple consequence of Chebyshev’s inequality that we have just derived is enough for our purposes.)

Sometimes we don’t know the characteristics of a probability space, and we want to estimate the mean of a random variable \( X \) by sampling its value repeatedly. (For example, we might want to know the average temperature at noon on a January day in San Francisco; or we may wish to know the mean life expectancy of insurance agents.) If we have obtained independent empirical observations \( X_1, X_2, \ldots, X_n \), we can guess that the true mean is approximately

\[ \hat{\mu} = \frac{X_1 + X_2 + \cdots + X_n}{n} \quad (8.19) \]