**Search on the web about the Law of large numbers LLN and compare it with Part b of your homework 3 and express in your own words whether your simulation is somehow related with this theorem, and why.**

Relation to Part B of the Homework: In Part B of the homework, I'm simulating the cumulated frequency of penetration (f) and examining how it relates to the number of attacks. As I collect more data points (simulated attacks), the frequency of penetration, when expressed as f/number of attacks or f/√number of attacks, becomes more stable and reliable.

The Law of Large Numbers is related to the simulation in the following ways:

Stability with More Data: As I simulate more attacks (increase the sample size), your calculated frequencies (f, f/number of attacks, f/√number of attacks) become more stable and accurate. This aligns with the LLN, which states that larger sample sizes lead to more accurate estimates.

Convergence to True Frequency: Just like LLN suggests that sample averages converge to the true population mean, in your simulation, the frequency of penetration (f) approaches the true probability of penetration as you consider a larger number of attacks. This convergence illustrates the principles of LLN.

Reduced Variability: With a larger number of attacks, the relative frequency (f/number of attacks) and the normalized ratio (f/√number of attacks) become less variable and closer to the true proportions. This reduction in variability is a key aspect of LLN, demonstrating that sample estimates become more reliable with increased sample size.

**Search on the web about the Central Limit Theorem CLT and compare it with Part a of your homework 3 and say in your own words whether your simulation is somehow related with this theorem, and why.  
Based on the CLT, how could you modify ("normalize") the "security score" to obtain an asymptotic convergence to a proper distribution?**

Part A of your homework involves simulating the accumulation of a "security score" for M systems subjected to N attacks. You're modeling this accumulation by assigning a score of -1 if a system is penetrated and 1 if it's successfully protected. You've assumed a constant penetration probability p for each attack. In this simulation, you're essentially tracking the sum of these scores over multiple attacks for each system.

The Central Limit Theorem (CLT) is a fundamental concept in statistics that states that the sum (or average) of many independent, identically distributed random variables will approximately follow a normal distribution, regardless of the underlying distribution of the variables themselves. This means that, under certain conditions, as you collect more and more data points and calculate the sum of these data points, the resulting distribution of the sums will resemble a bell-shaped normal distribution.

In the context of your Part A simulation, you are accumulating security scores over multiple attacks for each system. Each security score is an independent random variable with values of -1 and 1, based on the outcomes of the attacks. The CLT is relevant here because it suggests that as you accumulate more and more security scores for each system, the distribution of the total scores for all systems will tend towards a normal distribution.

To relate this to the CLT, imagine that you run your simulation for many systems (M) and a large number of attacks (N). Then, you calculate the total security scores for all systems for a particular attack number. By repeating this process for various attack numbers, you'll obtain a distribution of total scores. According to the CLT, as you increase M and N, this distribution of total scores will become more and more like a normal distribution.

Now, in Part B, you're asked to simulate the cumulated frequency of penetration, relative frequency (f/number of attacks), and the "normalized" ratio (f/√number of attacks). This is where the CLT can be applied. The CLT tells us that the relative frequency and the normalized ratio will also tend towards normal distributions as the number of attacks (N) becomes large, assuming that the penetration probability (p) is constant.

The CLT also implies that if you want to "normalize" the "security score" to obtain an asymptotic convergence to a proper distribution, you can consider the relative frequency or the "normalized" ratio you mentioned. These normalized values will be centered around a mean and will follow a normal distribution with a specific variance as you collect more data points (i.e., as the number of attacks increases).

In summary, the Central Limit Theorem is relevant to your simulation because it explains how the distribution of cumulative security scores (Part A) and the normalized values (Part B) will behave as you accumulate more data, specifically tending toward a normal distribution due to the independence and identical distribution of your security scores in each case.