The Central Limit Theorem (CLT)

The Central Limit Theorem (CLT) tells us that, under certain conditions, the result of adding together many random outcomes is approximately Gaussian distributed.

There are multiple versions of the theorem with different technical conditions and details. As is common, I'll sloppily refer to *the* CLT. The most important conditions and details, as you need to know it for this course, is outlined below. As we're not going to use a formal statement about the convergence to prove anything, I'm not going to be more precise.

Bounded mean and variance: The main requirement is that each variable included in the sum should have mean and variance below some fixed bound. Intuitively, if really extreme outcomes are common, they can dominate the sum, and the distribution of a single outcome can change the shape of the distribution of the sum away from a bell-curve.

Constrained values: If we add together integers, then the sum can only be an integer, so cannot be Gaussian distributed. In general, if some totals are impossible, the probability density function or mass function over the total value still tends to one that's proportional to a Gaussian PDF, but constrained to the possible values.

Convergence only close to the mean: Finally, the convergence guaranteed by theory is of a weak form, which only provides meaningful guarantees close to the mean. Only expect the PDF of the sum to be close to a Gaussian within a small number of standard deviations of the mean. The extreme tails of the distribution do not converge rapidly to a Gaussian. Don't assume a sum is Gaussian distributed, and then report statistical significance based on evaluating the Gaussian fit several standard deviations away from its mean!

Further reading could start at Wikipedia:
`https://en.wikipedia.org/wiki/Central_limit_theorem`

1 Check your understanding

Write a Matlab/Octave or Python program to demonstrate Central Limit behaviour. I say "program", but only a few lines of code are needed.

- Generate K random outcomes from some non-Gaussian distribution.
- Add them up, and record the total.
- Repeat the procedure N times to get N totals. Or do these operations all at once using matrix operations.

Compare the distribution of your N outcomes to a Gaussian. Can you make K big enough to get a reasonably Gaussian-like distribution? Does the K you need depend on the distribution of the terms in the sum?