

Questions for Discussion

1.1

Why do all the Gaussians in the first example have the same height, while the ones in the second example do not? Answer both in terms of the formula, and in plain English.

The "height" of a Gaussian is $\frac{1}{\sigma\sqrt{2\pi}}$. Because all the Gaussians in the first example had the same σ value, they had the same height. Likewise, because all the Gaussians in the second example had different σ values, they had different heights.

1.2

Specifying all three of x , μ and σ gives you a single value. Specifying only μ and σ gives you a curve as a function of x . In your own words, describe why sometimes we might want to be using the whole curve rather than a single value.

Sometimes, it'll be more useful to operate at a higher level of abstraction (i.e., that of random variables) and only materialize a probability at the very end.

2.1

Notice in the figure above that the μ and σ of the Gaussian are very close to the mean and standard deviation of the distribution of sums. How do you think these would change if we added together more numbers, say 24 numbers instead of 12 numbers?

The Irwin-Hall distribution (which describes the sum of uniform random variables) would predict that increasing the number of numbers would cause the resultant distribution to more closely approximate a normal distribution. With 24 numbers, the mean would be around 12 and the standard deviation would be around $\sqrt{2}$.

2.2

Notice in the figure above that the μ and σ of the Gaussian are very close to the mean and standard deviation of the distribution of n_{evt} . How do you think these would change if we set a higher average number of events per trial?

As we increase the λ parameter of the Poisson random variable, then μ and σ of the approximating normal random variable will also increase (because μ is approximately λ and σ is approximately $\sqrt{\lambda}$).

3.1

Different fields have different conventions for what constitutes a "statistically significant" result. In many fields a p-value of < 0.05 is considered significant. In other fields, the threshold is 5σ . How much more unlikely is a 5σ fluctuation than one with a p-value of 0.05? Why do you think that different fields might use such different conventions? (It is ok to guess...)

Around 174,428 times more unlikely! Practitioners in different fields may have different techniques and instruments available with varying degrees of accuracy and precision. Fields with more accurate/precise techniques/instruments probably adopt lower p-value thresholds than those with less accurate/precise techniques/instruments.