Physics 250
Approach to the central limit theorem.

Peter Young
(Dated: November 29, 2007)

Consider a random variable x with distribution $P(x)$. This has mean μ and standard deviation σ. According to the central limit theorem, if μ and σ are finite, the distribution of the sum of N independent such variables,

$$Y = \sum_{i=1}^{N} x_i \, ,$$

is, for $N \to \infty$, a Gaussian with mean $N\mu$ and standard deviation $\sqrt{N}\sigma$. It is convenient to subtract off the mean, and divide by $\sqrt{N}$, i.e. let

$$X = \frac{Y - N\mu}{\sqrt{N}} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (x_i - \mu) \tag{1}$$

because the central limit theorem then predicts that the distribution of X, which we call $P^{(N)}(X)$, becomes *independent of* N for large N, namely a Gaussian with zero mean and standard deviation unity:

$$\lim_{N\to\infty} P^{(N)}(X) = \frac{1}{\sqrt{2\pi}}\, e^{-X^2/2} \, . \tag{2}$$

Clearly $P^{(1)}(X) \equiv P(X + \mu)$, the distribution of the individual variables shifted so the mean is zero. We emphasize that even though $P(x)$ need not be a Gaussian, the distribution $P^{(N)}(X)$ will become Gaussian for large N (assuming the conditions of the central limit theorem hold; i.e. the mean μ and standard deviation of $P(x)$ are finite).

We illustrate the convergence to the central limit theorem as N is increased, by taking, for $P(x)$, the rectangular distribution

$$P(x) = \begin{cases} \dfrac{1}{2\sqrt{3}}, & (|x| < \sqrt{3}), \\ 0, & (|x| > \sqrt{3}) \, . \end{cases} \tag{3}$$

This is shown by the dotted line in Fig. 1, and is clearly quite different from a Gaussian, which is represented by the solid line. It is easy to see that

$$\mu \equiv \langle x \rangle = 0, \tag{4}$$

and a simple calculation gives

$$\sigma \equiv \left(\langle x^2 \rangle - \langle x \rangle^2 \right)^{1/2} = 1. \tag{5}$$

The distributions for $N = 2$ and 4 are shown by the short-dashed, and long-dashed lines in the figure. For $N = 2$, the distribution is a "tent" distribution (consisting of two straight lines; this can be shown analytically). It resembles a Gaussian more then the original rectangular distribution, but is not very close to it. However, we see that even for N as small as 4, the distribution $P^{(N)}(X)$ is very close to a Gaussian. For significantly larger values of N, the curves for $P^{(N)}(X)$ would be indistinguishable, in the figure, from the Gaussian curve.

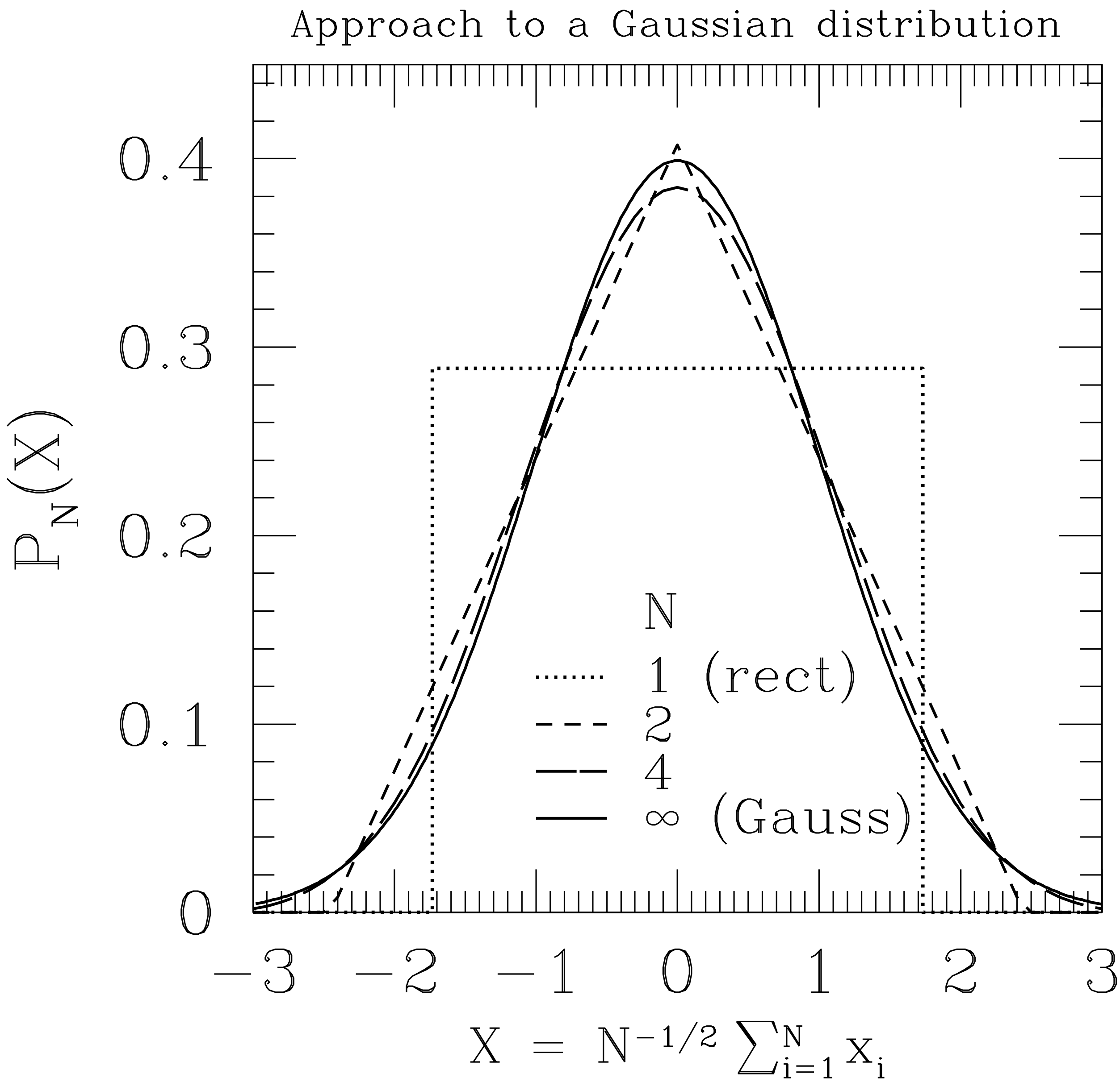

FIG. 1: The dotted line is the rectangular distribution, $P(X)$ ($\equiv P^{(1)}(X)$), in Eq. (3). The solid line is the Gaussian distribution in Eq. (2). The short-dashed and long-dashed lines are the distributions, $P^{(N)}(X)$, of the sum (divided by $\sqrt{N}$, see Eq. (1)) of $N = 2$ and 4 variables, each distributed according to the rectangular distribution.