will continue to be Gaussian. The fact that linear filter output to a Gaussian signal input will be a Gaussian signal is highly significant and is one of the most useful results in communication analysis.

6.7 CENTRAL LIMIT THEOREM

Under certain conditions, the sum of a large number of independent RVs tends to be a Gaussian random variable, independent of the probability densities of the variables added.* The rigorous statement of this tendency is what is known as the **central limit theorem**.† Proof of this theorem can be found in Refs. 6 and 7. We shall give here only a simple plausibility argument.

The tendency toward a Gaussian distribution when a large number of functions are convolved in shown in Fig. 6.20. For simplicity, we assume all PDFs to be identical, that is, a gate function $0.5\,\Pi(x/2)$. Figure 6.20 shows the successive convolutions of gate functions. The tendency toward a bell-shaped density is evident.

This important result that the **distribution** of the sum of n independent Bernoulli random variables, when properly normalized, converges toward Gaussian distribution was established first by A. de Moivre in the early 1700s. The more general proof for an arbitrary distribution was credited to J. W. Lindenber and P. Lévy in the 1920s. Note that the "normalized sum" is the sample average (or sample mean) of n random variables.

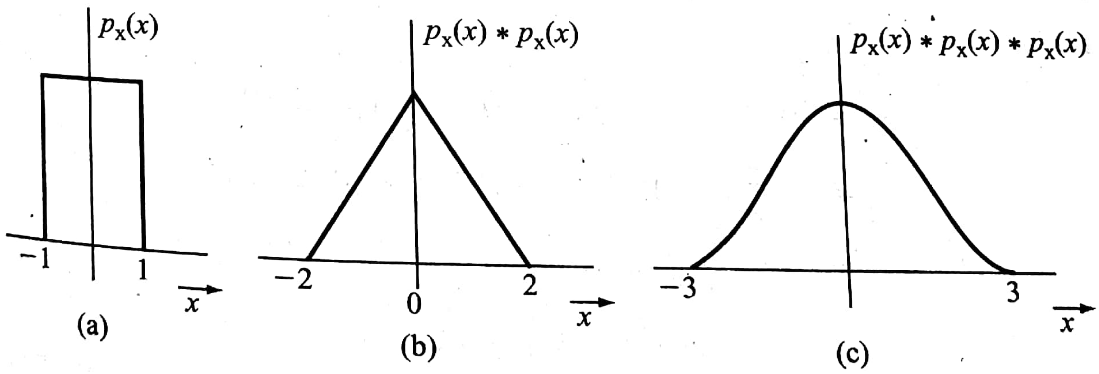

Figure 6.20 Demonstration of the central limit theorem.

Central Limit Theorem (for the sample mean):
Let $x_1, \ldots, x_n$ *be independent random samples from a given distribution with mean* μ *and variance* $0 < \sigma^2 < \infty$. *Then for any value x, we have*

$$\lim_{n\to\infty} P\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{x_i-\mu}{\sigma} \le x\right] = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\,dv \tag{6.99}$$

or equivalently,

$$\lim_{n\to\infty} P\left[\frac{\tilde{x}_n-\mu}{\sigma/\sqrt{n}} > x\right] = Q(x) \tag{6.100}$$

Note that

$$\tilde{x}_n = \frac{x_1+\cdots+x_n}{n}$$

is known as the sample mean. The interpretation is that the sample mean of any distribution with nonzero finite variance converges to Gaussian distribution with fixed mean μ *and decreasing variance* σ^2/n. *In other words, regardless of the true distribution of* x_i, $\sum_{i=1}^{n} x_i$ *can be approximated by a Gaussian distribution with mean* $n\mu$ *and variance* $n\sigma^2$.

Example 6.29

Consider a communication system that transmits a data packet of 1024 bits. Each bit can be in error with probability of 10^{-2}. Find the (approximate) probability that more than 30 of the 1024 bits are in error.

Solution

Define a random variable x_i such that $x_i = 1$ if the ith bit is in error and $x_i = 0$ if not. Hence

$$v = \sum_{i=1}^{1024} x_i$$

is the number of errors in the data packet. We would like to find $P(v > 30)$.

Since $P(x_i = 1) = 10^{-2}$ and $P(x_i = 0) = 1 - 10^{-2}$, strictly speaking we would need to find

$$P(v > 30) = \sum_{m=31}^{1024}\binom{1024}{m}\left(10^{-2}\right)^m\left(1-10^{-2}\right)^{1024-m}$$

This calculation is time-consuming. We now apply the central limit theorem to solve this problem approximately.

First, we find

$$\overline{x_i} = 10^{-2}\times(1) + (1-10^{-2})\times(0) = 10^{-2}$$
$$\overline{x_i^2} = 10^{-2}\times(1)^2 + (1-10^{-2})\times(0) = 10^{-2}$$

As a result,

$$\sigma_i^2 = \overline{x_i^2} - (\overline{x_i})^2 = 0.0099$$

Based on the central limit theorem, $v = \sum_i x_i$ is approximately Gaussian with mean of $1024\cdot 10^{-2} = 10.24$ and variance $1024 \times 0.0099 = 10.1376$. Since is a standard Gaussian with zero mean and unit variance,

$$P(v > 30) = P\left(y > \frac{30 - 10.24}{\sqrt{10.1376}}\right)$$
$$= P(y > 6.20611) = Q(6.20611)$$
$$\simeq 1.925 \times 10^{-10}$$

Now is a good time to further relax the conditions in the central limit theorem for the sample mean. This highly important generalization is proved by the famous Russian mathematician A. Lyapunov in 1901.

Central Limit Theorem (for the sum of independent random variables): *Let random variables* $x_1, \ldots, x_n$ *be independent but not necessarily identically distributed. Each of the random variable* x_i *has mean* μ_i *and nonzero variance* $\sigma_i^2 < \infty$*. Furthermore, suppose that each third-order central moment*

$$\overline{|x_i - \mu_i|^3} < \infty, \qquad i = 1, \ldots, n$$

and suppose

$$\lim_{n\to\infty} \sum_{i=1}^{n} \overline{|x_i - \mu_i|^3} \left(\sum_{i=1}^{n} \sigma_i^2\right)^{3/2} = 0$$

Then random variable

$$y(n) = \frac{\sum_{i=1}^n x_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

converges to a standard Gaussian density as $n \to \infty$, *that is,*

$$\lim_{n\to\infty} P\left[y(n) > x\right] = Q(x) \tag{6.101}$$

The central limit theorem provides a plausible explanation for the well-known fact that many rando variables in practical experiments are approximately Gaussian. For example, communication channel noise the sum effect of many different random disturbance sources (e.g., sparks, lightning, static electricity). Bas on the central limit theorem, noise as the sum of all these random disturbances should be approximat Gaussian.

6.8 FROM RANDOM VARIABLE TO RANDOM PROCESS

The notion of a random process is a natural extension of the random variable (RV). Consider, for exan the temperature x of a certain city at noon. The temperature x is an RV and takes on different values day. To get the ... values of x at noon over many days (a large ... temperature at noon)