Lecture 35

Andrei Antonenko

April 25, 2005

1 Central Limit Theorem

According to the weak law of large numbers, the distribution of the sample mean M_n is increasingly concentrated in the near vicinity of the true mean μ. In particular, its variance tends to zero. On the other hand, the variance of the sum $S_n = X_1 + \cdots + X_n = n \cdot M_n$ increases to infinity, and the distribution of S_n cannot be said to converge to anything meaningful. An intermediate view is obtained by considering the deviation $S_n - n\mu$ of S_n from its mean $n\mu$, and scaling it by a factor proportional to $1/\sqrt{n}$. What is special about this particular scaling is that it keeps the variance at a constant level. The central limit theorem asserts that the distribution of this scaled random variable approaches a normal distribution.

More specifically, let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with mean μ and variance σ^2. We define

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}. \tag{1}$$

An easy calculation yields

$$\mathbf{E}\left[Z_n\right] = \frac{\mathbf{E}\left[X_1 + \cdots + X_n\right] - n\mu}{\sigma\sqrt{n}} = 0, \tag{2}$$

and

$$\mathbf{var}\left(Z_n\right) = \frac{\mathbf{var}\left(X_1 + \cdots + X_n\right)}{\sigma^2 n} = \frac{\mathbf{var}\left(X_1\right) + \cdots + \mathbf{var}\left(X_n\right)}{\sigma^2 n} = \frac{n\sigma^2}{n\sigma^2} = 1. \tag{3}$$

Theorem 1.1 (Central Limit Theorem)**.** *Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with common mean μ and variance σ^2, and define*

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then, the CDF of Z_n converges to the standard normal CDF

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx,$$

in the sense that

$$\lim_{n\to\infty} P(Z_n \le z) = \Phi(z), \qquad \textit{for every } z.$$

The central limit theorem is surprisingly general. Besides independence, and the implicit assumption that the mean and variance are well-defined and finite, it places no other requirement on the distribution of the X_i, which could be discrete, continuous, or mixed random variables. It is of tremendous importance for several reasons, both conceptual, as well as practical. On the conceptual side, it indicates that the sum of a large number of independent random variables is approximately normal. As such, it applies to many situations in which a random effect is the sum of a large number of small but independent random factors. Noise in many natural or engineered systems has this property. In a wide array of contexts, it has been found empirically that the statistics of noise are well-described by normal distributions, and the central limit theorem provides a convincing explanation for this phenomenon. On the practical side, the central limit theorem eliminates the need for detailed probabilistic models and for tedious manipulations of PMFs and PDFs. Rather, it allows the calculation of certain probabilities by simply referring to the normal CDF table. Furthermore, these calculations only require the knowledge of means and variances.

2 Approximations Based on the Central Limit Theorem

The central limit theorem allows us to calculate probabilities related to Z_n as if Z_n were normal. Since normality is preserved under linear transformations, this is equivalent to treating S_n as a normal random variable with mean $n\mu$ and variance $n\sigma^2$.

Let $S_n = X_1 + \cdots + X_n$, where the X_i are independent identically distributed random variables with mean μ and variance σ^2. If n is large, the probability $P(S_n \leq c)$ can be approximated by treating S_n as if it were normal, according to the following procedure.

1. Calculate the mean $n\mu$ and the variance $n\sigma^2$ of S_n.

2. Calculate the normalized value
$$z = \frac{c - n\mu}{\sigma\sqrt{n}}. \tag{4}$$

3. Use the approximation
$$P(S_n \leq c) \approx \Phi(z), \tag{5}$$
where $\Phi(z)$ is available from standard normal CDF tables.

Example 2.1. We load on a plane 100 packages whose weights are independent random variables that are uniformly distributed between 5 and 50 pounds. What is the probability that the total weight will exceed 3000 pounds? It is not easy to calculate the CDF of the total weight and the desired probability, but an approximate answer can be quickly obtained using the central limit theorem. We want to calculate $P(S_{100} > 3000)$, where S_{100} is the sum of the 100 packages. The mean and the variance of the weight of a single package are

$$\mu = \frac{5 + 50}{2} = 27.5, \quad \sigma^2 = \frac{(50-5)^2}{12} = 168.75,$$

based on the formulas for the mean and variance of the uniform PDF. We thus calculate the normalized value

$$z = \frac{3000 - 100 \cdot 27.5}{\sqrt{168.75 \cdot 100}} = \frac{250}{129.9} = 1.92,$$

and use the standard normal tables to obtain the approximation

$$P(S_{100} \leq 3000) \approx \Phi(1.92) = 0.9726.$$

Thus the desired probability is

$$P(S_{100} > 3000) = 1 - P(S_{100} \leq 3000) \approx 1 - 0.9726 = 0.0274.$$

Example 2.2. A machine processes parts, one at a time. The processing times of different parts are independent random variables, uniformly distributed on $[1, 5]$. We wish to approximate the probability that the number of parts processed within 320 time units is at least 100. Let X_i be the processing time of the i-th part, and let $S_{100} = X_1 + \cdots + X_{100}$ be the total processing time of the first 100 parts. The event that the number of parts processed within 320 time units is at least 100 is the same as the event $\{S_{100} \leq 320\}$, and we can now use a normal approximation to the distribution of S_{100}. Note that

$$\mu = \mathbf{E}[X_i] = 3 \quad \text{and} \quad \sigma^2 = \mathbf{var}(X_i) = 16/12 = 4/3.$$

We calculate the normalized value

$$z = \frac{320 - n\mu}{\sigma\sqrt{n}} = \frac{320 - 300}{\sqrt{100 \cdot 4/3}} = 1.73,$$

and use the approximation

$$P(S_{100} \leq 320) \approx \Phi(1.73) = 0.9582.$$