and

$$s_n^2 = \sum_{k=1}^{n} \frac{n-k}{(n-k+1)^2} \sim \log n. \tag{6.6}$$

$S_n = X_1 + \ldots + X_n$ is the total number of cycles. *The average is m_n; the number of permutations with cycles between $\log n + \alpha(\log n)^{\frac{1}{2}}$ and $\log n + \beta(\log n)^{\frac{1}{2}}$ is given by $n!\{\Phi(\beta) - \Phi(\alpha)\}$, approximately.* The refined forms of the central limit theorem give more precise estimates.[11]

*7. THE STRONG LAW OF LARGE NUMBERS

The (weak) law of large numbers (5.4) asserts that for every particular sufficiently large n the deviation $|S_n - m_n|$ is likely to be small in comparison to n. It has been pointed out in connection with Bernoulli trials (chapter VIII) that this does not imply that $|S_n - m_n|/n$ remains small for all large n; it can happen that the law of large numbers applies but that $|S_n - m_n|/n$ continues to fluctuate between finite or infinite limits. The law of large numbers permits only the conclusion that large values of $|S_n - m_n|/n$ occur at infrequent moments.

We say that the sequence X_k obeys the strong law of large numbers if to every pair $\epsilon > 0$, $\delta > 0$, there corresponds an N such that there is probability $1 - \delta$ or better that for every $r > 0$ all $r + 1$ inequalities

$$\frac{|S_n - m_n|}{n} < \epsilon, \quad n = N, N+1, \ldots, N+r \tag{7.1}$$

will be satisfied.

We can interpret (7.1) roughly by saying that with an overwhelming probability $|S_n - m_n|/n$ *remains* small[12] for all $n > N$.

The Kolmogorov Criterion. *The convergence of the series*

$$\sum \frac{\sigma_k^2}{k^2} \tag{7.2}$$

[11] A great variety of asymptotic estimates in combinatorial analysis were derived by other methods by V. Gončarov, Du domaine d'analyse combinatoire, *Bulletin de l'Académie Sciences URSS, Sér. Math.* (in Russian, French summary), vol. 8 (1944), pp. 3–48. The present method is simpler but more restricted in scope; cf. W. Feller, The fundamental limit theorems in probability, *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 800–832.

* This section treats a special topic and may be omitted.

[12] The general theory introduces a sample space corresponding to the infinite sequence $\{X_k\}$. The strong law then states that with probability one $|S_n - m_n|/n$ tends to zero. In real variable terminology the strong law asserts convergence almost everywhere, and the weak law is equivalent to convergence in measure.

is a sufficient condition for the strong law of large numbers to apply to the sequence of mutually independent random variables X_k with variances σ_k^2.

Proof. Let A_ν be the event that for at least one n with $2^{\nu-1} < n \le 2^\nu$ the inequality (7.1) does *not* hold. Obviously it suffices to prove that for all ν sufficiently large ($\nu > \log N$) and all r

$$P\{A_\nu\} + P\{A_{\nu+1}\} + \ldots + P\{A_{\nu+r}\} < \delta,$$

that is, that the series $\Sigma P\{A_\nu\}$ converges. Now the event A_ν implies that for some n with $2^{\nu-1} < n \le 2^\nu$

$$|S_n - m_n| \ge \frac{\epsilon}{2} \cdot 2^\nu \tag{7.3}$$

and by Kolmogorov's inequality (chapter IX, section 7)

$$P\{A_\nu\} \le 4\epsilon^{-2} \cdot s_{2^\nu}^2 \cdot 2^{-2\nu}. \tag{7.4}$$

Hence

$$\sum_{\nu=1}^{\infty} P\{A_\nu\} \le 4\epsilon^{-2} \sum_{\nu=1}^{\infty} 2^{-2\nu} \sum_{k=1}^{2^\nu} \sigma_k^2 = 4\epsilon^{-2} \sum_{k=1}^{\infty} \sigma_k^2 \sum_{2^\nu \ge k} 2^{-2\nu} \le \le 8\epsilon^{-2} \sum_{k=1}^{\infty} \frac{\sigma_k^2}{k^2} \tag{7.5}$$

which accomplishes the proof.

As a typical application we prove the

Theorem. *If the mutually independent random variables X_k have a common distribution $\{f(x_j)\}$ and if $\mu = E(X_k)$ exists, then the strong law of large numbers applies to the sequence $\{X_k\}$.*

This theorem is, of course, stronger than the weak law of section 1. The two theorems are treated independently because of the methodological interest of the proofs. For a converse cf. problem 17.

Proof. We again use the method of truncation. Two new sequences of random variables are introduced by

$$\begin{aligned} U_k &= X_k, & V_k &= 0 & \text{if } |X_k| < k, \\ U_k &= 0, & V_k &= X_k & \text{if } |X_k| \ge k. \end{aligned} \tag{7.6}$$

The U_k are mutually independent, and we shall show that they satisfy Kolmogorov's criterion. Clearly

$$\sigma_k^2 \le E(U_k^2) = \sum_{|x_j| < k} x_j^2 f(x_j). \tag{7.7}$$