A Probabilistic and Combinatorial Analysis of a Digit–Sum Power Game

Fix integers $k \geq 2$ and $p \geq 1$. The game proceeds as follows. There are N tables, each with exactly m students, so that the total number of participants is $M = Nm$. Each student independently selects a k-digit base–10 integer n uniformly at random from the set $\{10^{k-1}, 10^{k-1}+1, \dots, 10^k - 1\}$. Equivalently, the leading digit is uniformly distributed on $\{1, \dots, 9\}$ and each of the remaining $k-1$ digits is uniformly distributed on $\{0, \dots, 9\}$, all digits being independent. Writing

$$n = \sum_{i=1}^{k} d_i 10^{k-i},$$

with $d_1 \in \{1, \dots, 9\}$ and $d_2, \dots, d_k \in \{0, \dots, 9\}$, define the digit sum

$$T = \sum_{i=1}^{k} d_i.$$

Each student computes the score

$$S = n - cT^p,$$

where the scalar $c \in \mathbb{R}$ is fixed in advance and announced before play begins. Each table computes the average of its students' scores, and the class outcome is defined as the average of the N table averages. Since all table sizes are equal, the class outcome coincides exactly with the grand mean

$$\overline{S}_M = \frac{1}{M} \sum_{j=1}^{M} S_j$$

of the M independent scores.

Under the uniform distribution on k-digit integers, the expectation of n is

$$\mathbb{E}[n] = \frac{10^{k-1} + (10^k - 1)}{2} = \frac{11 \cdot 10^{k-1} - 1}{2}.$$

The digit sum satisfies

$$\mathbb{E}[T] = 5 + 4.5(k-1) = 4.5k + 0.5, \qquad \mathrm{Var}(T) = \frac{20}{3} + (k-1)\frac{33}{4}.$$

The exact distribution of T is given combinatorially by the generating polynomial

$$P_k(x) = (x + x^2 + \cdots + x^9)(1 + x + \cdots + x^9)^{k-1},$$

so that

$$\mathbb{P}(T = t) = \frac{[x^t] P_k(x)}{9 \cdot 10^{k-1}},$$

and consequently

$$\mathbb{E}[T^p] = \frac{1}{9 \cdot 10^{k-1}} \sum_{t=1}^{9k} t^p [x^t] P_k(x).$$

For analytic purposes, a second–order Taylor expansion of x^p about $\mu_T = \mathbb{E}[T]$ yields the quadratic approximation

$$\mathbb{E}[T^p] \approx \mu_T^p + \tfrac{1}{2} p(p-1) \mu_T^{p-2} \sigma_T^2,$$

where $\sigma_T^2 = \mathrm{Var}(T)$.

To enforce unbiasedness, define

$$c = c^\star = \frac{\mathbb{E}[n]}{\mathbb{E}[T^p]},$$

so that $\mathbb{E}[S] = 0$ and hence $\mathbb{E}[\overline{S}_M] = 0$. Substituting the quadratic approximation gives

$$c^\star \approx \frac{\mathbb{E}[n]}{\mu_T^p + \frac{1}{2} p(p-1) \mu_T^{p-2} \sigma_T^2}.$$

The random variable S is bounded almost surely, since

$$10^{k-1} - c(9k)^p \le S \le 10^k - 1 - c.$$

In particular, all moments of S exist. Writing $\sigma_S^2 = \mathrm{Var}(S)$, one has the exact identity

$$\sigma_S^2 = \mathrm{Var}(n) + c^2 \mathrm{Var}(T^p) - 2c\, \mathrm{Cov}(n, T^p).$$

A quadratic approximation for σ_S^2 is obtained by expanding T^p to first order around μ_T, which gives

$$\mathrm{Var}(T^p) \approx p^2 \mu_T^{2p-2} \sigma_T^2, \qquad \mathrm{Cov}(n, T^p) \approx p \mu_T^{p-1} \mathrm{Cov}(n, T).$$

Since

$$\mathrm{Cov}(n, T) = \sum_{i=1}^{k} 10^{k-i} \mathrm{Var}(d_i) = 10^{k-1} \frac{20}{3} + \sum_{j=0}^{k-2} 10^j \frac{33}{4},$$

and

$$\mathrm{Var}(n) = \frac{10^{2k} - 10^{2k-2}}{12},$$

one obtains the quadratic variance approximation

$$\sigma_S^2 \approx \mathrm{Var}(n) + c^2 p^2 \mu_T^{2p-2} \sigma_T^2 - 2cp \mu_T^{p-1} \mathrm{Cov}(n, T),$$

with $c = c^\star$ or its quadratic approximation.

Because S is bounded, Hoeffding's inequality applies. For all $\varepsilon > 0$,

$$\mathbb{P}\big(|\overline{S}_M| \ge \varepsilon\big) \le 2 \exp\!\left(-\frac{2M\varepsilon^2}{(b-a)^2} \right),$$

where $a = 10^{k-1} - c(9k)^p$ and $b = 10^k - 1 - c$.

Moreover, since $\mathbb{E}[S] = 0$ and $0 < \sigma_S^2 < \infty$, the Berry–Esseen theorem implies that

$$\sup_x \left| \mathbb{P}\!\left(\frac{\sqrt{M}\, \overline{S}_M}{\sigma_S} \le x \right) - \Phi(x) \right| \le \frac{C\, \mathbb{E}|S|^3}{\sigma_S^3 \sqrt{M}},$$

where Φ denotes the standard normal distribution function and C is an absolute constant.

Consequently, for fixed $\alpha \in (0,1)$ and letting $z_{1-\alpha/2}$ denote the $(1-\alpha/2)$–quantile of the standard normal law, the central limit theorem yields the asymptotic design condition

$$\mathbb{P}\big(|\overline{S}_M| < \gamma\big) \approx 1-\alpha \quad \text{when} \quad M \approx \left(\frac{z_{1-\alpha/2}\sigma_S}{\gamma}\right)^2.$$