Let $y = (y_1, \ldots, y_k)$ be a random vector with the multinomial distribution, $y \sim \text{MN}_k(N, p)$, $p = (p_1, \ldots, p_k)$. The scaled deviance is defined as

$$d_k(y; p) := 2 \sum_{i=1}^k y_i \log \left( \frac{y_i}{Np_i} \right).$$

In a special case of the binomial deviance ($k = 2$), Hoeffding’s results [1] imply the tight exponential inequality:

$$\sup_{x, p_1} e^{x/2} \mathbb{P}\{d_2(y, p) \geq x\} = 2. \quad (1)$$

Here the supremum is taken over all $x > 0$, $p_1 \in (0, 1)$ and $N = 1, 2, \ldots$.

Using results of Zubkov and Serov [3], we show that there exists a function $G_N(x) = G_N(x | p_1)$ having closed-form representation and such that

$$\mathbb{P}\{d_2(y, p) \geq x\} \leq G_N(x | p_1), \quad \forall x > 0, \; p_1 \in (0, 1/2], \; N = 1, 2, \ldots, \quad (2)$$

and

$$e^{x/2} G_N(x | p_1) = O \left( N^{-1/2} \right) \quad (3)$$

uniformly with respect to $x \in [aN, bN]$ for some positive constants $a$ and $b > a$. The inequality (2) is locally-tight (up to a constant factor of 2) in the sense that, for all $p_1 \in (0, 1/2], \rho > \log(N) + 1 + \log((1 - p_1)/p_1)$, and $x \in (0, N|\log(p_1)|], N = 1, 2, \ldots,$

$$\sup_{t \in [x, x+\rho]} \frac{\mathbb{P}\{d_2(y, p) \geq t\}}{G_N(t | p_1)} \geq \frac{1}{2}. \quad (4)$$

We will discuss some extensions of (2)–(4) to $k > 2$ (cf. [2]).

References