M. V. Burnashev

ON THE DISTRIBUTION OF ONE STATISTICAL SUM FROM INFORMATION THEORY

Assume that $\boldsymbol{x}_j$, $j = 1, \ldots, M$, are independent and equiprobable binary n-vectors. Denote by $w_j = w(\boldsymbol{x}_j)$, $j = 1, \ldots, M$ - the weight (i.e. the number of ones) of binary n-vector $\{\boldsymbol{x}_j\}$. For $0 \le z \le 1$ consider the random sum

$$S(z, M, n) = \sum_{j=1}^{M} z^{w(\boldsymbol{x}_j)}. \tag{1}$$

The sum (1) arises often in information theory. Although all results below are non-asymptotic in n, M, they are mostly oriented to the case $n \to \infty$ and $M = e^{Rn}$, $R > 0$. Moreover, to simplify formulas we do not use below integer parts signs

Theorem 1. 1) For $0 \le z \le 1$ and $z^{n/2} \le A \le 1$ the following bounds hold

$$\begin{gathered} -\frac{\ln(n+1)}{n} \le \frac{1}{Mn} \ln \mathbf{P}\{S(z, M, n) \ge MA\} + \ln 2 - h(a_0) \le \frac{\ln n}{n}, \\ a_0 = \frac{\ln A}{n \ln z}, \qquad h(x) = -x \ln x - (1-x)\ln(1-x), \quad 0 \le a_0 \le 1/2. \end{gathered} \tag{2}$$

2) For $z^n \le A \le z^{n/2}$ the following bound holds

$$\frac{1}{Mn} \ln \mathbf{P}\{S(z, M, n) \le MA\} \le h(a_0) - \ln 2 + \frac{\ln(n+1)}{n}, \quad 1/2 \le a_0 \le 1. \tag{3}$$

2) For $z \ge 1$ and $1 \le A \le z^{n/2}$ the following bound holds

$$\frac{1}{Mn} \ln \mathbf{P}\{S(z, M, n) \le MA\} \le h(a_1) - \ln 2 + \frac{\ln(n+1)}{n}, \quad 1/2 \le a_1 \le 1. \tag{4}$$

In particular, we get from (2)-(3)

Corollary 1. For any $z > 0$ the following inequality holds

$$\mathbf{P}\left\{\left|\ln \frac{S(z, M, n)}{M z^{n/2}}\right| \ge \sqrt{n \ln(n+1)} |\ln z|\right\} \le (n+1)^{-M}. \tag{5}$$

Remark. It follows from (5) that if $M \sim e^{Rn}$, $R > 0$, then $S(z, M, n) \sim M z^{n/2}$ for any $z > 0$ with very high probability.