Capacity of the Binomial Channel, or Minimax Redundancy for Memoryless Sources

Christos Komninakis, Lieven Vandenberghe, and Richard D. Wesel
Department of Electrical Engineering, University of California, Los Angeles
e-mail: `{chkomn, wesel, vandenbe}@ee.ucla.edu`

Abstract— **This paper computes the capacity of the binomial channel of order n, for any finite n via convex optimization and duality. The binomial channel capacity is also the minimax coding redundancy for a class of memoryless sources. The exact capacity for finite n computed here is compared to the asymptotic expression found by Xie and Barron in [1].**

I. INTRODUCTION

Consider the "biased coin channel" or the binomial channel of order n. The input is a real random variable $\Theta \in [0,1]$, and the output is the result of n coin flips with probability of heads Θ. The number of heads $X \in \{k,\ k=0,1,\dots,n\}$ follows the binomial distribution with parameters n and θ:

$$t_k(\theta) \overset{\text{def}}{=} P(X=k \mid \theta) = \binom{n}{k}\theta^k(1-\theta)^{n-k}, k=0,\dots,n. \quad (1)$$

This paper presents an algorithm to determine the binomial channel capacity $C_n = \max_{dP(\theta)} I_n(\Theta;X)$ for any finite n, based on convexity and duality. This problem was solved asymptotically (as $n \to \infty$) in [1], treated from the dual viewpoint of universal encoding of a discrete memoryless Bernoulli source. The asymptotic minimax redundancy for universal encoding [1, p. 647] was computed:

$$\lim_{n\to\infty}\left(\max_{dP(\theta)}\int_0^1 D(\mathbf{t}(\theta)\parallel \mathbf{q})dP(\theta)\right) = \frac{1}{2}\log\frac{n}{2\pi e} + \log\pi, \quad (2)$$

where $\mathbf{q}$ is the pmf on X induced by the input distribution $dP(\theta)$ and the "channel law" $\mathbf{t}(\theta)$. Fig. 1 presents the binomial capacity for finite n along with the asymptotic result of [1].

Although the input alphabet of the binomial channel is uncountably infinite (the unit interval), the capacity achieving input distribution is concentrated at a finite number J of points $\theta_1,\ \theta_2,\ \dots\theta_J$ for any n, and $J \le n+1$. This was proven in [2] using Dubins' theorem (for a similar result see also Gallager's Cor. 4.5.3).

Despite the discreteness of the optimum "capacity achieving" input distribution, the well-known Blahut-Arimoto algorithm cannot be used to compute the capacity, because neither the input support $\{\theta_1,\ \theta_2,\ \dots\theta_J\}$ nor its cardinality J are known *a-priori*. Furthermore, $I_n(\Theta;X)$ is not convex in the variables $\theta_1,\ \dots\theta_J$. Our solution provides the optimum input support as a by-product: the support points $\theta_1,\ \dots\theta_J$ are located where the slack variables of the Lagrange dual problem become zero.

II. BINOMIAL CHANNEL CAPACITY

Consider $\mathcal{U}$ a σ-field on $[0,1]$ and $dF(\theta) \in \mathcal{F}$, the space of signed measures on $([0,1],\mathcal{U})$. Then the binomial channel capacity becomes a convex problem in Euclidean space:

$$\text{minimize } \sum_{k=0}^{n} q_k \log q_k + \int_0^1 h(\theta)dF(\theta) \quad (3)$$

subject to the inequality constraint $dF(\theta) \ge 0$ and the equality constraints $q_k = \int_0^1 t_k(\theta)dF(\theta)$, $k=0,1,\dots,n$, and $\int_0^1 dF(\theta) = 1$, where we define $h(\theta) = H(X|\Theta=\theta) = -\sum_{k=0}^{n} t_k(\theta)\log t_k(\theta)$.

Taking the Lagrange dual function and eliminating the slack variable of the inequality constraint, leads to the dual problem:

$$\min_{\mathbf{z}}\left\{\sum_{k=0}^{n}2^{-z_k} + \frac{e}{\log e}\cdot\max_{0\le\theta\le1}\left[\sum_{k=0}^{n} z_k t_k(\theta) - h(\theta)\right]\right\}. \quad (4)$$

This problem is finite-dimensional in the dual variable $\mathbf{z}$ in the Euclidean space $\mathbf{R}^{n+1}$, and may be solved via standard convex optimization techniques, such as the ellipsoid algorithm [3].

The optimal output pmf $\mathbf{q}^*$ is recovered from the dual optimal $\mathbf{z}^*$: $q_k^* = 1/e\cdot 2^{-z_k^*}, k=0,\dots,n$. The "complementary slackness" condition for inequality constraints [5, Thm. 2.1] (similar to [4, p. 218]) forces the input support to consist of exactly those θ_m where the inner function $\sum_{k=0}^{n} z_k t_k(\theta) - h(\theta)$ of (4) attains its maxima. By [2], there can only be $J \le n+1$ such θ_m, and usually there are significantly fewer. Since $\sum_{k=0}^{n} z_k^* t_k(\theta) - h(\theta) = -\log(e) + D(\mathbf{t}(\theta)\parallel\mathbf{q}^*)$, complementary slackness is another way of stating Csiszár's Min-Max Capacity Theorem [6], that $C = \min_{\mathbf{q}}\max_{dF(\theta)} D(\mathbf{t}(\theta)\parallel\mathbf{q})$. In other words, the optimum prior $dF^*(\theta)$ induces an output pmf $\mathbf{q}^*$ that is equidistant in the relative entropy sense from $\mathbf{t}(\theta)$ at those points θ where $D(\mathbf{t}(\theta)\parallel\mathbf{q}^*)$ is maximized.

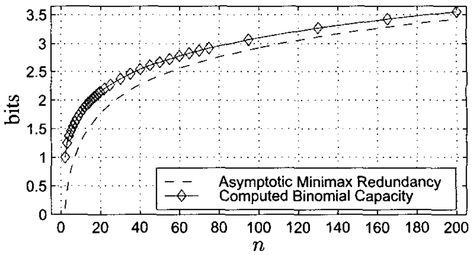

Figure 1: Asymptotic minimax redundancy in [1], and the binomial channel capacity for finite n.

The above method can also be applied to other problems with uncountably infinite input alphabet, where the capacity achieving prior distribution, as in the binomial channel, has finite support.

ACKNOWLEDGMENTS

The authors thank Tom Cover for identifying the capacity of the binomial channel as an open problem.

REFERENCES

[1] Q. Xie and A. Barron. Minimax redundancy for the class of memoryless sources. *IEEE Trans. on Info. Theory*, 43(2):646–657, March 1997.

[2] H.S. Witsenhausen. Some aspects of convexity useful in information theory. *IEEE Trans. on Info. Theory*, 26(3):265-71, May 1980.

[3] S. Boyd and L. Vandenberghe. Convex optimization. Course Reader for EE236B, Nonlinear Programming, UCLA, 1997.

[4] D. G. Luenberger. *Optimization by Vector Space Methods.* John Wiley & Sons, Inc., 1969.

[5] K. Isii. Inequalities of the types of Chebyshev and Cramèr-Rao and mathematical programming. *Annals of The Institute of Statistical Mathematics*, pages 277–293, 1964.

[6] I. Csiszár and J. Körner. *Information Theory: Coding theorems for discrete memoryless systems.* Academic Press, New York 1981.

0-7803-7123-2/01/$10.00 ©2001 IEEE