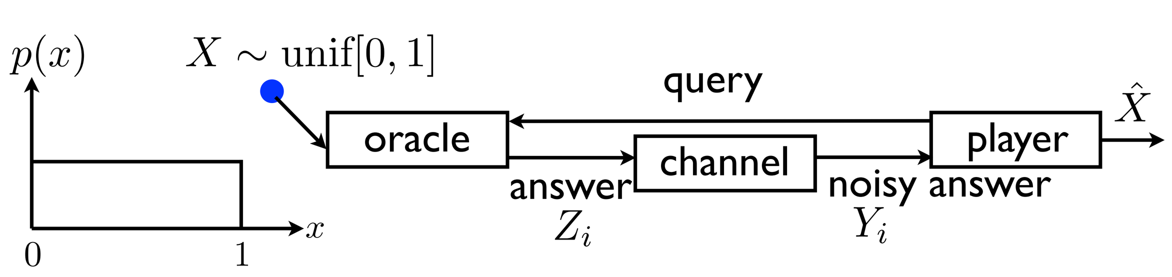
Bounds on Variance for Symmetric Unimodal Distributions

Hye Won Chung*, Brian M. Sadler† and Alfred O. Hero*
* University of Michigan, Ann Arbor, MI, USA
† US Army Research Laboratory, Adelphi, MD, USA

Introduction

Aim of the Project
Establish a direct relationship between differential entropy and variance for symmetric unimodal distributions.

Motivation

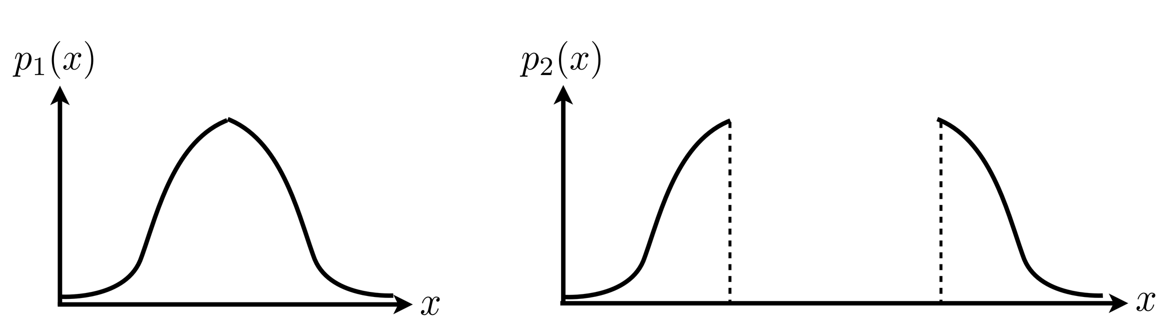


- Estimate the continuous random variable X by querying the oracle.
- Similar to the communication with noiseless feedback.
- Minimize the mean squared error $\mathbb{E}[|X-\hat{X}(Y_1^n)|^2]$.
- Q: What is the optimal querying strategy for this noisy case?
- Successive entropy minimization is often proposed as a way to progressively concentrate the posterior distribution.

Relationship Between Variance and Differential Entropy

$$\text{var}(X) = \int_{-\infty}^{\infty} (x-m)^2 p(x) dx,$$
$$h(p) = -\int_{-\infty}^{\infty} p(x)\log p(x) dx.$$

- No universal direct relation between two measures.
- The estimation counterpart to the Fano's inequality [1] shows $\frac{1}{2\pi e}e^{2h(p)} \leq \text{var}(X)$.
- In general, no upper bound on variance in terms of differential entropy.



$h(X) = -\int p_1(x)\log p_1(x)dx = -\int p_2(x)\log p_2(x)dx$

Previous Work: Log-Concave Distributions
In [2], it was shown that for a random variable X whose density $p(x)$ is log-concave, i.e.,

$$p(\alpha x + (1-\alpha)y) \geq p(x)^\alpha p(y)^{1-\alpha}$$

for each $x, y \in \mathbb{R}$ and each $0 \leq \alpha \leq 1$, the variance can be not only lower bounded but also upper bounded in terms of the entropy power, $e^{2h(p)}$, as

$$\frac{e^{2h(p)}}{2\pi e} \leq \text{var}(X) \leq c \cdot e^{2h(p)},$$

for some constant $c > 0$.

Main contributions

We establish the complementary result that such an upper bound on variance in terms of the entropy power, i.e.,

$$\frac{e^{2h(p)}}{2\pi e} \leq \text{var}(X) \leq c \cdot e^{2h(p)}, \quad c > 0$$

extends to the general class of symmetric unimodal distributions.

Symmetric unimodal mixtures

Symmetric Unimodal Mixture Densities

- Consider a linear mixture

$$p(x) = \sum_{i=1}^n \alpha_i p_i(x) \quad \text{for } \alpha_i \geq 0, \; \sum_{i=1}^n \alpha_i = 1, \quad (1)$$

  of either $p_i(x) \propto e^{-c_i|x-m|^{\theta_i}}$ for any $c_i > 0$, $\theta_i > 0$, or $p_i(x) = \text{unif}\left(-\frac{1}{2\epsilon_i}+m, \frac{1}{2\epsilon_i}+m\right)$ for any $\epsilon_i > 0$.
- Let σ_i^2 be the variance of $X \sim p_i(x)$.
- Define $r := \max_{i,j}\{\sigma_i^2/\sigma_j^2\}$.

Theorem 1 (Chung, Sadler, and Hero 2015)**.** *For a symmetric unimodal mixture density $p(x)$*

$$\frac{e^{2h(p)}}{2\pi e} \leq var(X) \leq cM(r)e^{2h(p)}$$

for some constant $c > 0$ and an increasing function $M(r) = \frac{(r-1)r^{\frac{1}{r-1}}}{e\log r} \geq 1$.

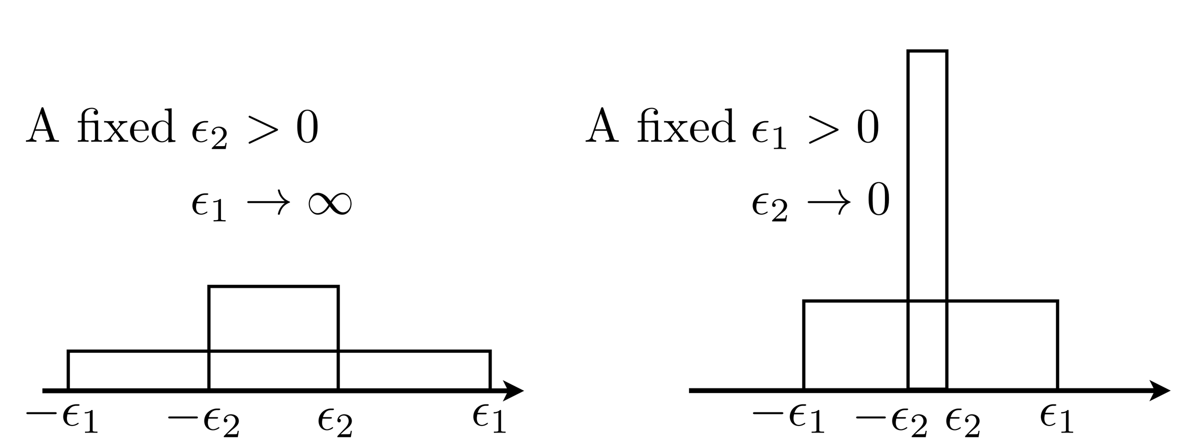
Sketch of Proof

1. For a mixture density $p(x) = \sum_{i=1}^n \alpha_i p_i(x)$, $\text{var}(X) = \sum_{i=1}^n \alpha_i\sigma_i^2$.
2. From the concavity of entropy, $e^{2h(p)} \geq \prod_{i=1}^n \left(e^{2h(p_i)}\right)^{\alpha_i}$.
3. From $\sigma_i^2 \propto e^{2h(p_i)}$, $e^{2h(p)} \geq c\left(\prod_{i=1}^n \left(\sigma_i^2\right)^{\alpha_i}\right)$ for some constant $c > 0$.
4. Reverse power mean inequality, $\sum_{i=1}^n \alpha_i\sigma_i^2 \leq M(r)\left(\prod_{i=1}^n \left(\sigma_i^2\right)^{\alpha_i}\right)$.
5. Therefore, $\text{var}(X) \leq c' M(r)e^{2h(p)}$ for some constant $c' > 0$.

Finite variance ratio

Necessity of Bounded Variance Ratio r
In Theorem 1, we assumed boundedness of the ratio r between the maximum and minimum variances of the mixture components. We show by a counterexample that this assumption is necessary.

Counterexample



- Consider $p(x) = \sum_{i=1}^2 \alpha_i \cdot \text{unif}\left(-\epsilon_i, \epsilon_i\right)$ for $\alpha_1, \alpha_2 > 0$, $\alpha_1 + \alpha_2 = 1$.
- The variance is proportional to the arithmetic mean, i.e., $\text{var}(X) \propto \left(\alpha_1\epsilon_1^2 + \alpha_2\epsilon_2^2\right)$.
- When $\frac{\epsilon_1}{\epsilon_2} \to \infty$, the entropy power is proportional to the geometric mean, i.e., $e^{2h(p)} \propto \epsilon_1^{2\alpha_1}\epsilon_2^{2\alpha_2}$.
- If $\epsilon_1 \to \infty$ for a fixed ϵ_2, variance increases much faster;
  If $\epsilon_2 \to 0$ for a fixed ϵ_1, $e^{2h(p)} \to 0$ but $\text{var}(X) > 0$.
- Boundedness of r is necessary for existence of an upper bound on variance in form of $ce^{2h(p)}$.

Lipschitz continuous densities

Lipschitz Continuous Symmetric Unimodal Density with Bounded Support

- Suppose that a symmetric unimodal density p with support $[m-s, m+s]$ satisfies the Lipschitz condition with constant $c_l \geq 0$, i.e.,

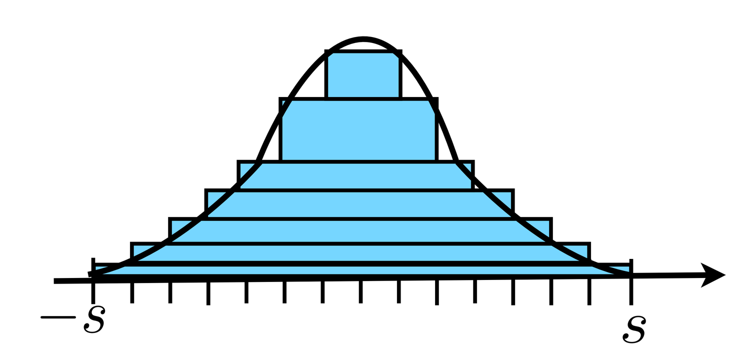
$$|p(x+y) - p(x)| \leq c_l|y|$$

  for any $x, y \in [m-s, m+s]$.
- The Lipschitz constant c_l and the support size s determines the tightness of the upper bound.

Theorem 2 (Chung, Sadler, and Hero 2015)**.** *For any Lipschitz continuous symmetric unimodal density $p(x)$ with bounded support $[m-s, m+s]$,*

$$\frac{e^{2h(p)}}{2\pi e} \leq var(X) \leq \frac{c_l s^2 e^{c_l s^2}}{24}e^{2h(p)}.$$

Sketch of Proof: Approximation of Symmetric Unimodal Distribution



- Construct a linear mixture density $\bar{p}_n(x) = \sum_{t=1}^n \alpha_t \cdot \text{unif}(-th, th)$ that approximates $p(x)$ such that

$$e^{2h(\bar{p}_n)} \leq e^{2h(p)}(1 + c_1 n^{-1}\log n);$$
$$|\text{var}(p) - \text{var}(\bar{p}_n)| \leq c_2 n^{-1}$$

  for some constants $c_1, c_2 > 0$.
- Prove an upper bound on variance of $\bar{p}_n(x)$ in terms of $e^{2h(\bar{p}_n)}$:

$$\text{var}(\bar{p}_n) \leq \frac{c_l s^2 e^{c_l s^2}}{24}e^{2h(\bar{p}_n)}\left(1 + c_3 n^{-1}\right).$$

- By combining these results and letting $n \to \infty$,

$$\text{var}(p) \leq \frac{c_l s^2 e^{c_l s^2}}{24}e^{2h(p)}.$$

Summary

Conditional Entropy as a Surrogate for MSE

- The variance of the general class of symmetric unimodal distribution can be bounded below and above by a constant scaling of its entropy $e^{2h(p)}$

$$\frac{e^{2h(p)}}{2\pi e} \leq \text{var}(X) \leq ce^{2h(p)}.$$

- Provide justification to successive entropy minimization in Bayesian sequential query design when the posterior is symmetric and unimodal.
- Future work: can we remove symmetry condition? extension to multi-dimensional case?

Acknowledgement: This research was partially supported by MURI grant W911NF-11-1-0391.

Reference:

[1] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
[2] S. Bobkov and M. Madiman, "The entropy per coordinate of a random vector is highly constrained under convexity conditions," *Information Theory, IEEE Transactions on*, vol. 57, no. 8, pp. 4940–4954, 2011.
[3] H. W. Chung, B. Sadler, and A. Hero, "Bounds on Variance for Symmetric Unimodal Distributions." Communication, Control, and Computing (Allerton), 2011 53th *Annual Allerton Conference on*. IEEE, 2015.