Consistent Estimation of Conditional Cumulants in the Empirical Bayes Framework (Extended Abstract)

Tang Liu*†, Alex Dytso*, H. Vincent Poor†, and Shlomo Shamai (Shitz)‡
*Department of Electrical and Computer Engineering, New Jersey Institute of Technology
†Department of Electrical and Computer Engineering, Princeton University
‡Department of Electrical Engineering, Technion – Israel Institute of Technology
Email: tang.liu@princeton.edu*†, alex.dytso@njit.edu*, poor@princeton.edu† and sshlomo@ee.technion.ac.il‡

2022 56th Asilomar Conference on Signals, Systems, and Computers | 978-1-6654-5906-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IEEECONF56349.2022.10052066

Background: Consider a noisy observation $Y = X + N$ where X is a random variable, and N is a Gaussian random variable with zero mean, variance σ^2, independent from X. The object of this work is to construct a consistent estimator for the conditional *cumulants* of the random variable X given the observation $Y = y$, in the empirical Bayes framework. Cumulants are important statistical quantities that provide useful alternatives to moments and have a variety of applications [1]–[4]. Given the conditional cumulant generating function

$$K_X(t|Y=y) = \log\left(\mathbb{E}[e^{tX}|Y=y]\right), y \in \mathbb{R}, t \in \mathbb{R}, \quad (1)$$

the conditional cumulant of order $k \in \mathbb{N}$ (set of non-negative integers) is defined as

$$\kappa_{X|Y=y}(k) = \frac{\mathrm{d}^k}{\mathrm{d}t^k} K_X(t|Y=y)|_{t=0}, y \in \mathbb{R}. \quad (2)$$

Concretely, in this work we are interested in the scenario where we observe n independent and identically distributed (i.i.d.) copies $Y_1, \ldots, Y_n$ of a random variable Y and seek to construct an estimator of $\kappa_{X|Y=y}(k)$ for all y and a given order k. Importantly, since neither the distribution nor observations of the random variable X are available, this puts us in the empirical Bayes framework [5].

Tools: The construction of the *consistent* estimator will rely on the following tools. The first tools is Tweedie's formula [5], which connects the conditional expectation $\mathbb{E}[X|Y=y]$ with the marginal pdf f_Y of Y

$$\mathbb{E}[X|Y=y] = y + \sigma^2 \frac{f'_Y(y)}{f_Y(y)}, y \in \mathbb{R}. \quad (3)$$

Note that Tweedie's formula allows one to compute $\mathbb{E}[X|Y=y]$ only from the knowledge of the marginal pdf f_Y.

The second tool that we will use is the fundamental relationship between conditional cumulants and the conditional expectation established in [6]

$$\kappa_{X|Y=y}(k+1) = \sigma^{2k} \frac{\mathrm{d}^k}{\mathrm{d}y^k} \mathbb{E}[X|Y=y],\ y \in \mathbb{R},\ k \in \mathbb{N}. \quad (4)$$

This tool will allow us to estimate cumulants providing that we have a good estimate of the conditional expectation. As the third tool, we will rely on *Lanczos' generalized derivatives* [7]–[10] defined as follows. Let $f : \mathbb{R} \to \mathbb{R}$ and define the following operator: for $h > 0$

$$D_h^{(n)} f(x) = \frac{c_n}{h^n} \int_{-1}^{1} f(x+ht) P_n(t) \mathrm{d}t,\ x \in \mathbb{R}, \quad (5)$$

where $c_n = \frac{1}{2}\sqrt{\frac{2^{2n+2}}{\pi}}\Gamma\left(n+\frac{3}{2}\right)$, and P_n is the *Legendre polynomial* of order n. The operator in (5) has the following property:

$$D_h^{(n)} f(x) = f^{(n)}(x) + O(h^2),\ n \in \mathbb{N}. \quad (6)$$

This property is useful for showing that empirical estimators of conditional cumulants are consistent.

Contributions: Inspired by (4) and (6), we consider the following estimator of $\kappa_{X|Y=y}$: for some $h > 0$

$$\widehat{\kappa}_{X|Y=y}(k+1; h) = \sigma^{2k} D_h^{(k)} \widehat{m}(y),\ y \in \mathbb{R},\ k \in \mathbb{N}, \quad (7)$$

where the estimator $\widehat{m}(y)$ of $\mathbb{E}[X|Y=y]$ is inspired by (8) and is given by

$$\widehat{m}(y) = y + \sigma^2 \frac{\hat{f}'_Y(y)}{\hat{f}_Y(y)}, y \in \mathbb{R}, \quad (8)$$

where $\hat{f}_Y(y) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{a} k\left(\frac{y-Y_i}{a}\right)$ is the kernel density estimator of $f_Y(y)$ with bandwidth $a > 0$ and kernel $k(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}$. Moreover, $\hat{f}'_Y(y)$ is the first order derivative of $\hat{f}_Y(y)$. The main result of this work is the following which shows that the estimator in (7) is consistent.

Theorem 1. *Let $t_n = \frac{\sigma^2\sqrt{w\log(n)}}{3}$, $a = \frac{1}{n^u}$ and $w \in (0, u)$ for some $u \in \left(0, \frac{1}{8}\right)$. Moreover, assume that $\mathbb{E}[X^2] < \infty$. Then, for every $k \in \mathbb{N}_{>0}$ and $\sigma^2 > 0$*

$$\lim_{n\to\infty} \mathbb{P}\left[\max_{|y| \le \frac{t_n}{2}} \left|\kappa_{X|Y=y}(k) - \widehat{\kappa}_{X|Y=y}(k; h_n)\right| \ge \frac{\tilde{C}_{k,\sigma}\, t_n}{n^{\frac{2(u-w)}{1+k}}}\right] = 0,$$

where $h_n = \left(\frac{C_\sigma}{n^{u-w}}\right)^{\frac{1}{k+1}}$ and C_σ and $\tilde{C}_{k,\sigma}$ are constants that depend only on k and σ.

Acknowledgements: The work of H. V. Poor has been supported by the U.S. National Science Foundation under Grant CCF-1908308. The work of S.S. (Shitz) has been supported by the US-Israel Binational Science Foundation (BSF) under grant BSF-2018710.

REFERENCES

[1] T. A. Courtade and S. Verdú, "Cumulant generating function of codeword lengths in optimal lossless compression," *IEEE International Symposium on Information Theory*, 2014.

[2] L. D. Lathauwer, J. Castaing, and J.-F. Cardoso, "Fourth-order cumulant-based blind identification of underdetermined mixtures," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2965 – 2973, 2007.

[3] P. Comon, "Independent component analysis, a new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[4] M. Zidane and R. Dinis, "Modeling and identification of nonlinear effects in massive mimo systems using a fifth-order cumulants-based blind approach," *Applied Sciences*, vol. 12, no. 7, 2022.

[5] H. Robbins, "An empirical Bayes approach to statistics," in *Proc. Third Berkeley Symp. Math Statist. Probab.* Citeseer, 1956.

[6] A. Dytso, H. V. Poor, and S. Shamai, "A general derivative identity for the conditional mean estimator in gaussian noise and some applications," *arXiv:2104.01883*, 2021.

[7] C. Lanczos, *Applied Analysis*. Prentice Hall, 1956.

[8] C. Groetsch, "Lanczos' generalized derivative," *The American Mathematical Monthly*, vol. 105, no. 4, pp. 320–326, 1998.

[9] S. Rangarajan and S. P. Purushothaman, "Lanczos' generalized derivative for higher orders," *J. Comput Appl. Math*, vol. 177, no. 2, pp. 461–465, 2005.

[10] J. Shen, "On the generalized 'lanczos' generalized derivative," *The American Mathematical Monthly*, vol. 106, no. 8, pp. 766–768, 1999.