A Note on the Distribution of Weighted Rademacher Random Variables

Ohad Asor*

December 10, 2014

Abstract

We prove some new results regarding inner product with ± 1 RVs, giving exact characterization of Rademacher random variables. As an application we present a method of solving the Partition problem.

We discuss random variables of the form $\langle \sigma, \mathbf{x} \rangle$ (where $\langle \cdot, \cdot \rangle$ is the usual $L_2 [\mathbb{R}^n]$ inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{n} x_n y_n$) over uniform $\sigma \in \{-1, 1\}^n$. We give an explicit form of the characteristic function and cumulant generating function, including all cumulants explicity. We also formulate the probaility mass function for the integer case expressed as an integral. We begin with a lemma taken from Wikipedia, then immediately state our main theorem:

Lemma 1. *For all* $\mathbf{x} \in \mathbb{R}^n$*:*

$$\prod_{k=1}^{n} \cos x_k = \mathbb{E}_\sigma \cos \langle \sigma, \mathbf{x} \rangle$$

Proof. (of Lemma 1) For $n = 1$ we have $\cos x = \frac{1}{2} (\cos (x_k) + \cos (-x_k))$ which is true under the cosine being an even function. Assuming correctness for n, then

$$\prod_{k=1}^{n+1} \cos x_k = \cos x_{n+1} \prod_{k=1}^{n} \cos x_k = 2^{-n} \sum_{\sigma} \cos x_{n+1} \cos \left(\sum_{k=1}^{n} \sigma_k x_k \right)$$

Since $\cos x \cos y = \frac{1}{2} (\cos (x + y) + \cos (x - y))$:

$$2^{-n-1} \sum_{\sigma} \cos \left(\sum_{k=1}^{n} \sigma_k x_k + x_{n+1} \right) + \cos \left(\sum_{k=1}^{n} \sigma_k x_k - x_{n+1} \right) = 2^{-n-1} \sum_{\sigma} \cos \left(\sum_{k=1}^{n+1} \sigma_k x_k \right)$$

$\square$

*ohadasor@gmail.com

Theorem 2. *Let $\mathbf{x} \in \mathbb{R}^n$ and denote by $X = \langle \sigma, \mathbf{x} \rangle$ a random variable depending on uniform random $\sigma \in \{-1, 1\}^n$. Then:*

1. The characteristic function of X is

$$\psi(t) = \prod_{k=1}^{n} \cos t x_k$$

2. Denote[1] $K_{2m} = 0, K_{2m+1} = B_{2m} 4^m (4^m - 1)$. The cumulants of X are

$$\kappa_{2m} = K_{2m-1} \sum_{k=1}^{n} x_k^{2m}$$

3. If $\mathbf{x} \in \mathbb{Z}^n, b \in \mathbb{Z}$, then:

$$\mathcal{P}[X = b] = \frac{1}{\pi (1 + \delta_b^0)} \int_{-\pi}^{\pi} \cos(bt) \prod_{k=1}^{n} \cos(x_k t)\, dt$$

   where the Delta is Kronecker's.

4. If $\mathbf{x} \in \mathbb{Z}^n$, then:

$$\mathcal{P}[X = 0] = 1 + \sum_{m=2}^{\infty} (-1)^m \frac{\mu_m}{(2m+1)!}$$

   where $\mu_{2m+1} = 0$ and $\mu_{2m} = \kappa_{2m} + \sum_{t=2}^{2m-1} \binom{2m-1}{2t-1} \kappa_{2t} \mu_{2m-2t}$.

Proof. Using Lemma 1, the linearity of expectation and the symmetry yielding $\mathbb{E}_\sigma [f(\sigma)] = \mathbb{E}_\sigma [f(-\sigma)]$ for any f:

$$\psi(t) = \prod_{k=1}^{n} \cos t x_k = \mathbb{E}_\sigma \cos \langle \sigma, t\mathbf{x} \rangle = \frac{1}{2} \mathbb{E}_\sigma \left[e^{it\langle \sigma, \mathbf{x} \rangle} + e^{-it\langle \sigma, \mathbf{x} \rangle} \right] = \mathbb{E}_\sigma \left[e^{it\langle \sigma, \mathbf{x} \rangle} \right]$$

and 1 is shown. Our CGF is therefore:

$$\ln \prod_{k=1}^{n} \cos(-itx_k) = \sum_{k=1}^{n} \ln \cosh(tx_k) = \sum_{k=1}^{n} x_k \int \tanh(x_k t)\, dt$$

Consequently, 2 is shown by:

$$\kappa_m = \frac{\partial^m}{\partial t^m} \left(\sum_{k=1}^{n} x_k \int \tanh(x_k t)\, dt \right) \Bigg|_{t=0}$$

[1] B_n are Bernoulli numbers.

To derive 3, recall any integers c, d satisfy $\int_{-\pi}^{\pi} \cos(cx)\cos(dx)\,dx = \pi\delta_c^d\left(1+\delta_c^0\right)$. Now:

$$\frac{1}{\pi\left(1+\delta_b^0\right)}\int_{-\pi}^{\pi}\psi(t)\cos(bt)\,dt = \frac{1}{\pi\left(1+\delta_b^0\right)}\mathbb{E}_\sigma\int_{-\pi}^{\pi}\cos(t\langle\sigma,\mathbf{x}\rangle)\cos(bt)\,dt = \mathbb{E}_\sigma\delta^b_{\langle\sigma,\mathbf{x}\rangle} = \mathcal{P}\left[\langle\sigma,\mathbf{x}\rangle = b\right]$$

Recalling that sinc function returns zero for all integers except zero, and the fact its Taylor series converges over the whole real line:

$$\mathcal{P}\left[\langle\sigma,\mathbf{x}\rangle = 0\right] = \mathbb{E}_\sigma\left[\frac{\sin\langle\sigma,\mathbf{x}\rangle}{\langle\sigma,\mathbf{x}\rangle}\right] = \sum_{m=0}^{\infty}(-1)^m\frac{\mathbb{E}_\sigma\langle\sigma,\mathbf{x}\rangle^{2m}}{(2m+1)!}$$

and the moments are computed from the cumulants. In fact, sinh can be used as well in a similar form:

$$\mathcal{P}\left[\langle\sigma,\mathbf{x}\rangle = 0\right] = \mathbb{E}_\sigma\left[\frac{\sinh\langle\sigma,\mathbf{x}\rangle}{\langle\sigma,\mathbf{x}\rangle}\right] = \sum_{m=0}^{\infty}\frac{\mathbb{E}_\sigma\langle\sigma,\mathbf{x}\rangle^{2m}}{(2m+1)!}$$

□

Remark 3. Since $\mathbf{x}^T\Sigma\mathbf{y} = \left\langle \mathrm{vec}(\Sigma), \mathrm{vec}\left(\mathbf{x}\mathbf{y}^T\right)\right\rangle$, our results can be extended to $\{\mathbf{x},\mathbf{y}\}\subset\mathbb{R}^n$ and Σ is $\{-1,1\}^{n\times n}$ matrix.

Example 4. The Partition problem may be written as: "Given $\mathbf{x}\in\mathbb{Z}^n$, does exist $\sigma\in\{-1,1\}^n$ such that $\langle\mathbf{x},\sigma\rangle = 0$?" which can be resolved by the series expansion of $\mathcal{P}[X=0]$ up to accuracy 2^{-n}.

Problem 5. What is the characteristic function of the r.v. $\langle\mathbf{x},\sigma\rangle\cdot\langle\mathbf{y},\sigma\rangle$?

Acknowledgement. Special credit to the pseudonym HunterMinerCrafter which contributed endless days verifying the results theoretically and empirically, while suggesting various tactics and way of thinking.

References

[1] van Zuijlen Martien CA (2011) On a conjecture concerning the sum of independent Rademacher random variables. http://arxiv.org/abs/1112.4988

[2] Yu, J. (2004). "Empirical characteristic function estimation and its applications". Econometrics Reviews 23 (2): 93–1223. doi:10.1081/ETC-120039605.

[3] Cormen, Thomas H.; Leiserson, Charles E., Rivest, Ronald L., Stein, Clifford (2001) [1990]. "35.5: The subset-sum problem". Introduction to Algorithms (2nd ed.). MIT Press and McGraw-Hill. ISBN 0-262-03293-7.

[4]