Proofs for the Maximum Entropy Property of the Normal Distribution

Christoph F. Mecklenbräuker

Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology, Vienna, Austria

September 19, 2010

It is well known that for any absolutely continuous random variable, the distribution that maximizes the differential entropy subject to an upper bound σ^2 on its second moment is the zero-mean normal distribution with variance σ^2.
In this contribution, several proofs for the maximum entropy property of the normal distribution are reviewed: Calculus of variations [Shannon 48, Kapur 89], use of Jensen's inequality [McEliece 77], and exploitation of the information inequality [Cover and Thomas 91], as well as Gallager's proof [Gallager 68]. The discussion emphasizes the corresponding concepts and pedagogical aspects.

References

[Shannon 48] C.E. Shannon, "A Mathematical Theory of Communication", Bell Syst. Tech. J., Vol. 27, July, pp. 379–423; Oct., pp. 623–656, 1948.

[Gallager 68] R.G. Gallager, "Information Theory and Reliable Communication", Wiley, New York, 1968.

[McEliece 77] R.J. McEliece, "The Theory of Information and Coding: A Mathematical Framework for Communication," Addison Wesley, Reading, Mass., 1977.

[Kapur 89] J.N. Kapur: Maximum-entropy models in science and engineering, John Wiley and Sons, New York, 1989.

[Cover and Thomas 91] T. Cover and J.A. Thomas, "Elements of Information Theory", John Wiley and Sons, New York, 1991.

[Gibson 93] J.D. Gibson, "Principles of Digital and Analog Communications," 2^{nd} ed., Maxmillan Publishing Co., New York, 1993.