Differential Entropy - Seminar 7

November 26, 2013

Problem 1 (Differential Entropy) *Evaluate the differential entropy $h(X) = -\int f \ln f$ for the following:*

(a) The exponential density, $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.

(b) The Laplace density, $f(x) = \frac{1}{2}\lambda \exp(-\lambda |x|)$, $x \in \mathbb{R}$.

(c) The sum of two independent normal RVs with means μ_i and variances σ_i^2, $i = 1, 2$.

Solution.

(a) Exponential

$$
\begin{aligned}
h(f) &= -\int_0^\infty \lambda e^{-\lambda x} \ln\left(\lambda e^{-\lambda x}\right) \mathrm{d}\, x = -\int_0^\infty \lambda e^{-\lambda x} \left[\ln \lambda - \lambda x\right] \mathrm{d}\, x \\
&= -\ln \lambda + 1 \text{ nats} \\
&= \log \frac{e}{\lambda} \text{ bits}
\end{aligned}
$$

(b) Laplace

$$
\begin{aligned}
h(f) &= -\int_{-\infty}^\infty \frac{1}{2}\lambda \exp(-\lambda |x|) \left(\ln \frac{1}{2} + \ln \lambda - \lambda |x|\right) \mathrm{d}\, x \\
&= -2\int_0^\infty \frac{1}{2}\lambda \exp(-\lambda x) \left(\ln \frac{1}{2} + \ln \lambda - \lambda x\right) \mathrm{d}\, x \\
&= -\left(\ln \frac{1}{2} + \ln \lambda\right) \int_0^\infty \lambda \exp(-\lambda x) \mathrm{d}\, x + \lambda^2 \int_0^\infty x \exp(-\lambda x) \mathrm{d}\, x \\
&= -\ln \lambda - \ln \frac{1}{2} + 1 = \ln \frac{2e}{\lambda} \text{ nats} \\
&= \log \frac{2e}{\lambda} \text{ bits.}
\end{aligned}
$$

(c) If $X \sim N(\mu, \sigma^2)$ then $h(X) = \frac{1}{2} \log 2\pi e \sigma^2$. Since $Y = X_1 + X_2 \sim N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$

$$
h(Y) = \frac{1}{2} \log 2\pi e \left(\sigma_1^2 + \sigma_2^2\right).
$$

■

Problem 2 (Concavity of determinants) *Let K_1 and K_2 be two symmetric nonnegative definite $n \times n$ matrices. Prove the result of Ky Fan[1]:*

$$\left|\lambda K_1 + (1-\lambda) K_2\right| \geq \left|K_1\right|^{\lambda} \left|K_2\right|^{1-\lambda} \qquad \text{for } 0 \leq \lambda \leq 1,$$

*where $|K|$ denotes the determinant of K. [*Hint*: Let $Z = X_\theta$, unde $X_1 \sim N(0, K_1)$, $X_2 \sim N(0, K_2)$ and $\theta = Bernoulli(\lambda)$. Then use $h(Z|\theta) \leq h(Z)$.]*

Solution. Let the RV θ havet he distribution of $P(\theta = 1) = \lambda$, $P(\theta = 2) = 1 - \lambda, 0 \leq \lambda \leq 1$. Let θ, X_1, X_2 independent and $Z = X_\theta$. Then Z has the covariance $K_Z = \lambda K_1 + (1-\lambda)K_2$. Z will not be multivariate normal. However, since a normal distribution maximizes the entropy for a given variance, we have

$$\begin{aligned}\frac{1}{2}\ln\left(2\pi e\right)^n \left|\lambda K_1 + (1-\lambda)K_2\right| &\geq H(Z) \geq h(Z|\theta) \\ &= \frac{1}{2}\lambda \ln\left(2\pi e\right)^n \left|K_1\right| + \frac{1}{2}(1-\lambda)\ln\left(2\pi e\right)^n \left|K_2\right|.\end{aligned}$$

Thus

$$\left|\lambda K_1 + (1-\lambda) K_2\right| \geq \left|K_1\right|^{\lambda} \left|K_2\right|^{1-\lambda},$$

as desired. ■

Problem 3 (Quantized random variables) *Roughly how many bits are required on the average to describe to three-digit accuracy the decay time (in years) of a radium atom if the half-life of radium is 80 years? Note that half-life is the median of the distribution and the distribution of the process is exponential.*

Solution. The differential entropy of an exponential RV with mean $\frac{1}{\lambda}$ is $\log \frac{e}{\lambda}$ bits. If the median is 80 years, then

$$\int_0^{80} \lambda e^{-\lambda x} \mathrm{d}\, x = \frac{1}{2}$$

or $1 - e^{-80\lambda} = \frac{1}{2}$. Solution is: $\left\{\frac{1}{80}\ln 2 - \frac{1}{40} i\pi l \mid l \in \mathbb{Z}\right\}$; the real solution is $\frac{1}{80}\ln 2 = 8.664\,3 \times 10^{-3}$. To represent the RV with 3 digits of accuracy ≈ 10 bits we need $\log_2 \frac{e}{8.664\,3\times 10^{-3}} + 10 = 18.293$ bits. ■

Problem 4 (Scaling) *Let $h(X) = -\int f(x)\log f(x)\mathrm{d}\, x$. Show $h(AX) = \log|det(A)| + h(X)$.*

Solution. Let $Y = Ax$. Then the density of Y is

$$g(y) = \frac{1}{|A|} f\left(A^{-1}y\right).$$

Hence

$$
\begin{aligned}
h(AX) &= -\int g(y) \ln g(y) \mathrm{d}\, y \\
&= -\int \frac{1}{|A|} f\left(A^{-1} y\right)\left[\ln f\left(A^{-1} y\right) - \log |A|\right] \mathrm{d}\, y \\
&= \int \frac{1}{|A|} f(x) \left[\ln f(x) - \log |A|\right] |A| \,\mathrm{d}\, x \qquad (1) \\
&= h(X) + \log |A| . \qquad (2)
\end{aligned}
$$

■

Problem 5 (Shape of the typical set) *Let X_i be i.i.d. $\sim f(x)$, where*

$$f(x) = ce^{-x^4}.$$

Let $h = -\int f \ln f$. Describe the shape (or form) or the typical set

$$A_{\varepsilon}^{(n)} = \left\{ x^n \in \mathbb{R}^n : f(x^n) \in \left(2^{-n(h+\varepsilon)}, 2^{-n(h-\varepsilon)}\right)\right\}.$$

Solution. Since the X_i are i.i.d., $f(x^n) = c^n e^{-\left(x_1^4 + \cdots + x_n^4\right)}$. From the definition of the typical set, x^n is typical if and only if

$$2^{-n(h+\varepsilon)} \leq c^n e^{-\left(x_1^4 + \cdots + x_n^4\right)} \leq 2^{-n(h-\varepsilon)}$$

which is equivalent to

$$-n(h+\varepsilon)\ln 2 \leq n \ln c - \left(x_1^4 + \cdots + x_n^4\right) \leq -n(h-\varepsilon)\ln 2$$

Finally, re-arranging the preceding condition, we have

$$A_{\varepsilon}^{(n)} = \left\{ x^n \in \mathbb{R}^n : n(\ln C + (h-\varepsilon)\ln 2) \leq \sum x_i^4 \leq n(\ln C + (h+\varepsilon)\ln 2)\right\}.$$

■

References

[1] Ky Fan, On a theorem of Weyl concerning the eigenvalues of linear transformations II, *Proc. Nat. Acad. Sci*, **36**: 31-35, 1950.