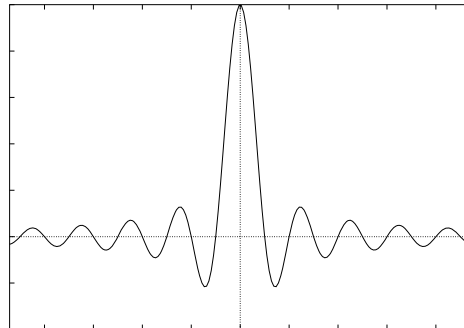


9 Information Theory (jgd1000)

- (a) A machine learning system was trained to learn about random variable  $X$  using a training set in which the probability distribution of its values  $\{x_i\}$  was  $p(x)$ , having entropy  $H(p)$ ; but then a test data set had a different distribution  $q(x)$ .
- (i) Define the cross-entropy  $H(p, q)$  between distributions  $p(x)$  and  $q(x)$ . What condition minimises its value, and then what value is it? [3 marks]
- (ii) Define the Kullback-Leibler distance (or the relative entropy)  $D_{KL}(p||q)$  between the distributions  $p(x)$  and  $q(x)$ . What condition minimises its value, and then what value is it? [3 marks]
- (iii) Show that:  $H(p, q) = D_{KL}(p||q) + H(p)$ . [2 marks]
- (b) The function  $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$  for  $x \neq 0$  as plotted below plays an important role in the Sampling Theorem. By considering its Fourier transform, show that this function is unchanged in form after convolution with itself, and show that it even remains unchanged in form after convolution with any higher frequency sinc function,  $\text{sinc}(ax)$  for  $a > 1$ ; but that if  $0 < a < 1$ , then the result of the convolution is instead that lower frequency sinc function  $\text{sinc}(ax)$ . [4 marks]



- (c) Explain the use of run-length encoding in JPEG compression. How does it enable compression factors of typically at least 10:1 with no perceptible loss, and even compression factors of 30:1 or higher? What is the role of the quantisation table in achieving this, and what is the relevant fact about spatial frequency sensitivity in human visual perception? [3 marks]
- (d) (i) Define the *genetic isopoint* of a human population. (ii) For most Europeans today, in what century did it occur? (iii) For a large well-mixed population of size  $m$ , approximately how many generations  $N$  ago was the genetic isopoint? (iv) Regarding genetic transmission as a lossy information channel, what sampling fact becomes critical for the effect of an ancestor once a family tree extends back at least  $N = 15$  generations? (v) What does Information Theory imply is achieved by sexual (as opposed to asexual) reproduction? [5 marks]