PD Dr. Torsten Enßlin
Vishal Johnson, Ming Huang

Exercise sheet 2

Exercise 2 - 1

Assume that loosing a fraction x of your budget hits you as $l(x) = x/(1-x)$, being your personal loss function.

a) Up to which fraction $y < x < 1$ of your budget should you invest to insure against the risk of losing the budget fraction $x = x_E$ by an event E occurring with probability $p = P(E)$ (2 points) ?

b) An insurance company asks you to pay a budget fraction $z = \alpha p x$ to insure your loss under E with $\alpha > 1$ (to ensure that the company makes profit on average). Under which conditions should you take their offer (1 point)?

c) Show that a tiny monetary loss $x \ll 1$ should never be insured (1 point).

d) The insurance company knows the true p_E perfectly well, whereas the beliefs of its customers on it are uniformly distributed in $p \in [0, 1]$ What is the expected profit per potential costumer as a function of the insurance price z (2 points) ?

Exercise 2 - 2

Imagine you wanted to store a probability distribution $P(x)$ over two events $x \in \{0, 1\}$ on a computer with very limited memory and precision. Of course it is enough to store only $P(0)$ since $P(1)$ can then be calculated through normalization, but you still have to round the numbers up to machine precision. Let X be the set of all numbers the computer can represent. Furthermore, let

$$q_{\text{low}} = \max\{q | q \in X \wedge q \leq P(0)\}$$
$$\text{and } q_{\text{high}} = \min\{q | q \in X \wedge q \geq P(0)\}$$

i.e., q_{low} the highest number in X that is still lower than $P(0)$ and q_{high} the lowest number in X that is still higher than $P(0)$.

a) Derive a decision rule for when to round to q_{low} or q_{high} for general $P(0)$, q_{low} and q_{high} based on the rule that you want to lose the least amount of information from the original distribution $P(x)$. (2 points)

b) Using the decision rule derived in a), determine wether it is better to round

- $P(1) = 0.146$ to 0.1 or 0.2
- $P(1) = 0.01$ to 0 or 0.5?

(1 point)

Exercise 2 - 3

A sequence of n coin tosses is performed and stored in a data vector $d^{(n)} = (d_1, \ldots d_n) \in \{0, 1\}^n$. The coin produced a head (denoted by a 1 in the data vector) with constant, but unknown frequency $f = P(d_i = 1|f) \in [0, 1]$.

a) How many bits of extra information on f are provided by the data vector $d^{(n)} = (1, \ldots 1)$ of only ones?

Hint: The extra information contained in a probability distribution $p(x)$ compared to a probability distribution $q(x)$ (in bits) is given by $\int \mathrm{d}x\, p(x) \log_2 \left(\frac{p(x)}{q(x)}\right)$. Furthermore you may use the following integral formulas:
$\int_0^1 dx\, x^n (1-x)^m = \frac{n!\,m!}{(n+m+1)!}$ for $n, m \in \mathbb{N}$, $\int_0^1 dx\, x^n \ln x = -\frac{1}{(n+1)^2}$ (2 points)

b) After how many such sequential heads did one obtain 10 bits of information on f? An accuracy of 10% is sufficient.

Hint: If $n > 10$ you can use $\frac{n}{n+1} \approx 1$. Use $2^{1/\ln 2} = 2^{\ln e/\ln 2} = 2^{\log_2 e} = e \approx 2.7$. (1 point)

c) How many bits on the outcome of the next toss is provided by a sequence of n heads? Provide also the asymptotic for $n \to \infty$!

Hint: It is helpful to guess the maximal amount of obtainable information before the detailed calculation is done. (2 points)