We first define the function h and put the number of digits high, to have better computer precision

> $h(x) := -\dfrac{(x \cdot \log(x) + (1-x) \cdot \log(1-x))}{\log(2)};$

> $Digits := 250;$

We compute the values for b,a and c which are related with the atomic distribution which shows that the improvement is bounded (by c).

> $with(RootFinding):$
$B := NextZero\left(x \mapsto h(x) * (2 - h(x)) - h\left(2 * x - x^2\right), 0.14\right);$

> $evalf[10](B);$

$$0.3294547385 \tag{1}$$

> $A := \dfrac{(1 - h(B))}{(2 - h(B))};$

> $evalf[10](A);$

$$0.07887729268 \tag{2}$$

> $C := A + (1 - A) \cdot B;$

> $evalf[10](C);$

$$0.3823455334 \tag{3}$$

Next, we consider the ratio's of the derivatives to compute the optimal choice of alpha.

> $g1(x) := diff\left((1 - (C - x) / (1 - x))^{\wedge}2 * h\left(2 * x - x^2\right) - (1 - (C - x) / (1 - x)) * h(x), x\right);$

> $g2(x) := diff((1 - 2 * (C - x) / (1 - x)) - (1 - (C - x) / (1 - x)) * h(x), x);$

> $\mathrm{alpha} := -\dfrac{eval(g1(x), x = B)}{(eval(g2(x), x = B) - eval(g1(x), x = B))};$

> $evalf[10](\mathrm{alpha})$

$$0.03560698066 \tag{4}$$

We verify the statement for probability distributions for which the support contains 3 elements; a_1,a_2 and 1, which have probabilities respectively equal to p_1, p_2 and 1-p_1-p_2. As such we can compute E [H(p)], E[H(p+q-pq)] and E'[H(min(2*p,1/2)]

>

> $H(p\_1, a\_1, p\_2, a\_2) := p\_1 \cdot h(a\_1) + p\_2 \cdot h(a\_2);$

> $Hpq(p\_1, a\_1, p\_2, a\_2) := p\_1^2 \cdot h\left((1 - a\_1)^2\right) + p\_2^2 \cdot h\left((1 - a\_2)^2\right) + 2 \cdot p\_1 \cdot p\_2 \cdot h((1 - a\_1) \cdot (1 - a\_2));$

> $Hpr(p\_1, a\_1, p\_2, a\_2) := (p\_1 + \min((p\_2 - (1 - p\_1 - p\_2)), 0)) \cdot h\left(\min\left(2 \cdot a\_1, \frac{1}{2}\right)\right) + \max((p\_2 - (1 - p\_1 - p\_2)), 0) \cdot h\left(\min\left(2 \cdot a\_2, \frac{1}{2}\right)\right);$

We do the optimization problem, once in the precise sense. Here we conclude that the extremal distribution seems to be atomic.

> $with(Optimization):$
$Minimize((1 - \mathrm{alpha}) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \mathrm{alpha} \cdot Hpr(p\_1, a\_1, p\_2, a\_2) - H(p\_1, a\_1, p\_2, a\_2), \{p\_1 \le 1, p\_2 \le 1, a\_1 \le C, a\_2 \le 0.5, p\_1 \cdot a\_1 + p\_2 \cdot a\_2 + (1 - p\_1$

$- p\_2) \leq C, p\_1 + p\_2 \leq 1, a\_1 \leq a\_2\}, assume = nonnegative)$

Once, we add the condition the probability p_1 has to be strictly positive and a_1 is bounded by e. g. 0.26, the minimization problem has a strict positive output. From this, we conclude that the extremum indeed seems to occur for the atomic distribution.

> $with(Optimization):$
> $Minimize((1 - \text{alpha}) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \text{alpha} \cdot Hpr(p\_1, a\_1, p\_2, a\_2) - H(p\_1, a\_1, p\_2, a\_2), \{0.0000001 \leq p\_1, p\_1 \leq 1, a\_1 \leq 0.26, p\_2 \leq 1, a\_2 \leq 0.5, p\_1 \cdot a\_1 + p\_2 \cdot a\_2 + (1 - p\_1 - p\_2) \leq C, p\_1 + p\_2 \leq 1, a\_1 \leq a\_2\}, assume = nonnegative)$
> ;

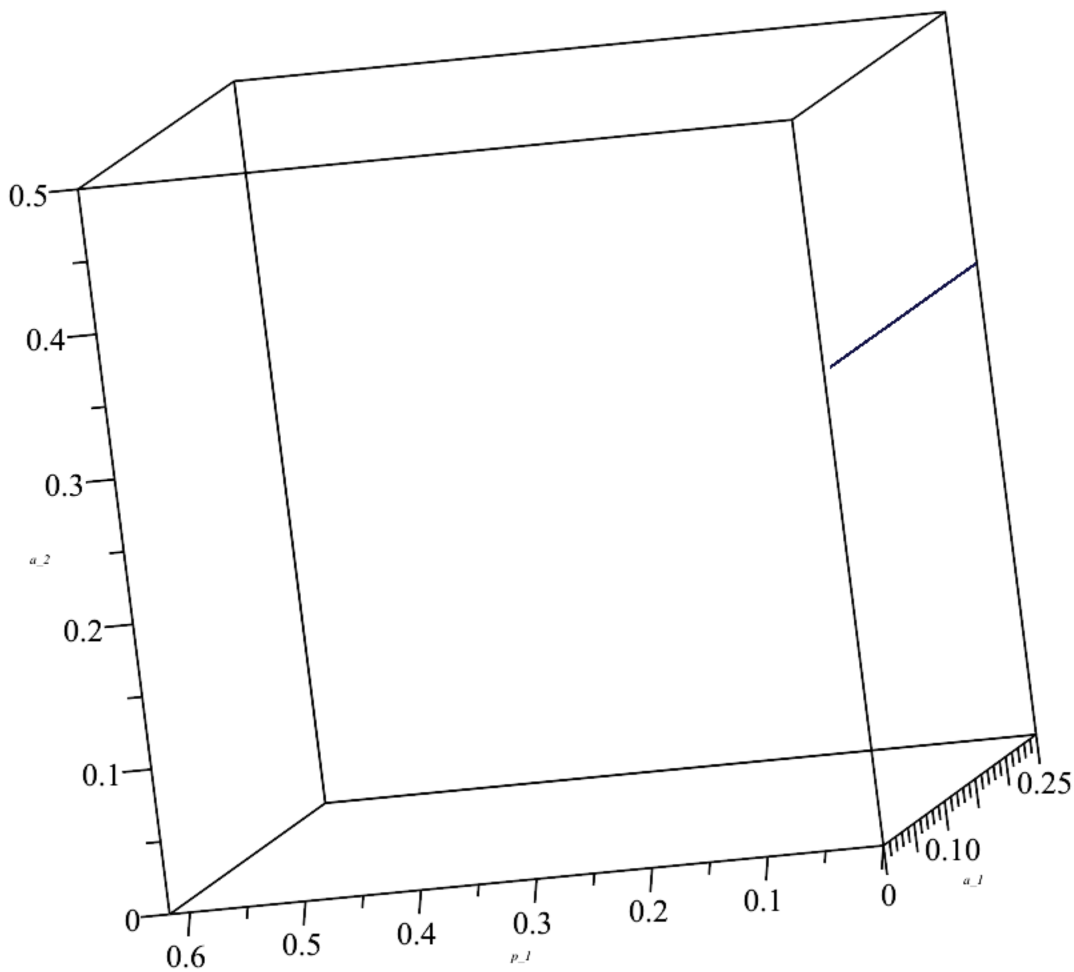
>

> $F(p\_1, a\_1, p\_2, a\_2) := (1 - \text{alpha}) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \text{alpha} \cdot Hpr(p\_1, a\_1, p\_2, a\_2) - H(p\_1, a\_1, p\_2, a\_2)$

$$F := (p\_1, a\_1, p\_2, a\_2) \mapsto (1 - \alpha) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \alpha \cdot Hpr(p\_1, a\_1, p\_2, a\_2) - H(p\_1, a\_1, p\_2, a\_2) \quad (5)$$

>

> $plots\text{:-}implicitplot3d\left(F\left(p\_1, a\_1, \frac{(p\_1 \cdot a\_1 + 1 - p\_1 - C)}{1 - a\_2}, a\_2\right) - 0.00001, a\_1 = 0 .. 0.25, p\_1 = 0 ..1 - C, a\_2 = 0 .. 0.5, numpoints = 10000, style = surface, color = navy\right);$



(6)