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

> $h(x) := -\frac{(x \cdot \log(x) + (1-x) \cdot \log(1-x))}{\log(2)};$
$h(0) := 0;$
$h(1) := 0;$

> $Digits := 250;$

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).

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

> $evalf[10](B);$

$$0.3294547385 \tag{1}$$

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

> $evalf[10](A);$

$$0.07887729268 \tag{2}$$

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

> $evalf[10](C);$

$$0.3823455334 \tag{3}$$

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

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

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

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

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

$$0.03560698066 \tag{4}$$

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 − alpha)·Hpq(p_1, a_1, p_2, a_2) + alpha· Hpr(p_1, a_1, p_2, a_2) −H(p_1,
      a_1, p_2, a_2), {p_1 ≤ 1, p_2 ≤ 1, a_1 ≤ C, a_2 ≤ 0.5, p_1·a_1 + p_2·a_2 + (1 − p_1
      − p_2) ≤ C, p_1 + p_2 ≤ 1, a_1 ≤ 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 − alpha)·Hpq(p_1, a_1, p_2, a_2) + alpha· Hpr(p_1, a_1, p_2, a_2) −H(p_1,
      a_1, p_2, a_2), { 0 ≤ p_1, p_1 ≤ 1, a_1 ≤ 0.26, p_2 ≤ 1, a_2 ≤ 0.5, p_1·a_1 + p_2·a_2
      + (1 − p_1 − p_2) ≤ C, p_1 + p_2 ≤ 1, a_1 ≤ a_2}, assume = nonnegative) ;
[−2.80 × 10^(−248), [a_1                                                                  (5)
    = 0.24576492240787172144531099604567114334318102052334087545756212861630587 94\
    0576787167859646869425280631524094125506683067749086103748131288969352867980 9\
    6837361167896150008370295445369461912957859897758252494970149726700159948388 7\
    758074966928360876489 75, a_2
    = 0.32945473850303697239170538387713413075183911691669851293736770887488124 07\
    0884012308855656570127619887302146407104407813684067828083220680568730224567 6\
    2611508602298093032526948077583092934755173983122470927912920510207860026231 4\
    1999124570859772313066 9, p_1 = 0., p_2
    = 0.92112270729407682658783805546470314439041491005542401413080045579046746 74\
    5182768285932527532179576165984941650384213028727697024854129963868820302822 1\
    9365488315263598229721877143978669824206030894672790940214945829663171178430 5\
    787772699703495507711 01]]
```

```
> with(Optimization) :
  Minimize( (1 − alpha)·Hpq(p_1, a_1, p_2, a_2) + alpha· Hpr(p_1, a_1, p_2, a_2) −H(p_1,
      a_1, p_2, a_2), { 0.1 ≤ p_1, 10^(−100) ≤ a_1, p_1 ≤ 1, a_1 ≤ 0.1, p_2 ≤ 1, a_2 ≤ 0.5, p_1
      ·a_1 + p_2·a_2 + (1 − p_1 − p_2) ≤ C, p_1 + p_2 ≤ 1, a_1 ≤ a_2}, assume
      = nonnegative) ;
[                                                                                         (6)
    4.5244321054442659125108016759988173107910988793377497543967139397374984874 81\
    0943426499682405783078075686884221367794126098148045419844216338467408573164 8\
    5910078902504721071544336470342871203916365752077396617930624090484887403046 2\
    4612565911282796567 × 10^(−99), [a_1
    = 1.000000000000000000000000000000000000000000000000000000000000000000000000\
    00000000000000000000000000000000000000000000000000000000000000000000000000000\
    00 × 10^(−100), a_2
    = 0.01271577109808289991778354952781759048789410738688175835914852595080135 04\
    3213560097428915447873013627284490641317155745381472945200526565020993791379 7\
    2700942513100678255138661892712150644118692353451439633371652238101847692044 1\
```

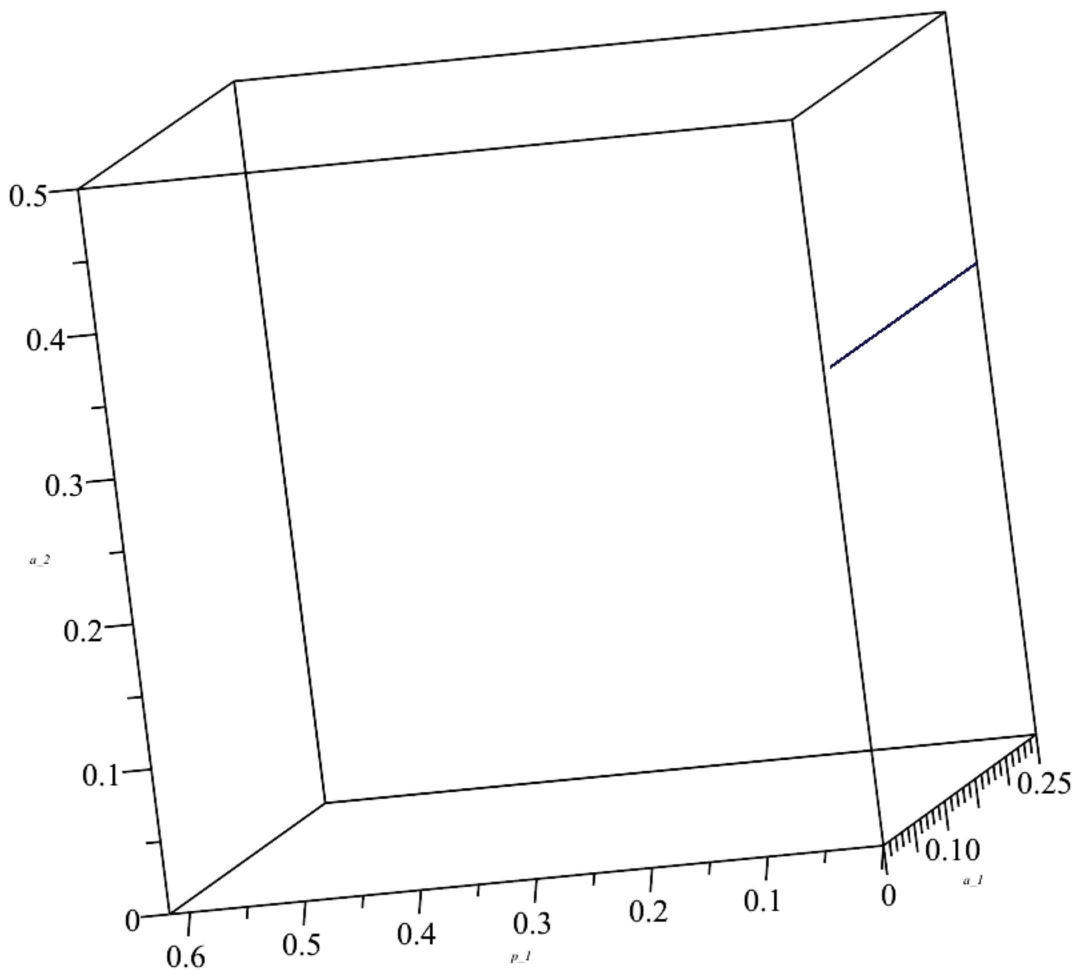
617681544778013432965838, p_1
$= 0.619584995743932143503745543455980443763163030807472225110411601949948981\backslash$
$8562825552150578906279184710880514217046077390913419558344365309254768823043326\backslash$
$382984909127246941798094267406472865380510020728904576770308429498020816734033\backslash$
$22292756081016899164 41, p\_2 = 0.]]$

```
> F(p_1, a_1, p_2, a_2) := (1 - alpha)·Hpq(p_1, a_1, p_2, a_2) + alpha· Hpr(p_1, a_1, p_2, a_2) -H(p_1, a_1, p_2, a_2)
```

$$F := (p\_1, a\_1, p\_2, a\_2) \mapsto (1 - \alpha) \cdot Hpq(p\_1, a\_1, p\_2, a\_2) + \alpha \cdot Hpr(p\_1, a\_1, p\_2, a\_2) - H(p\_1, a\_1, p\_2, a\_2) \tag{7}$$

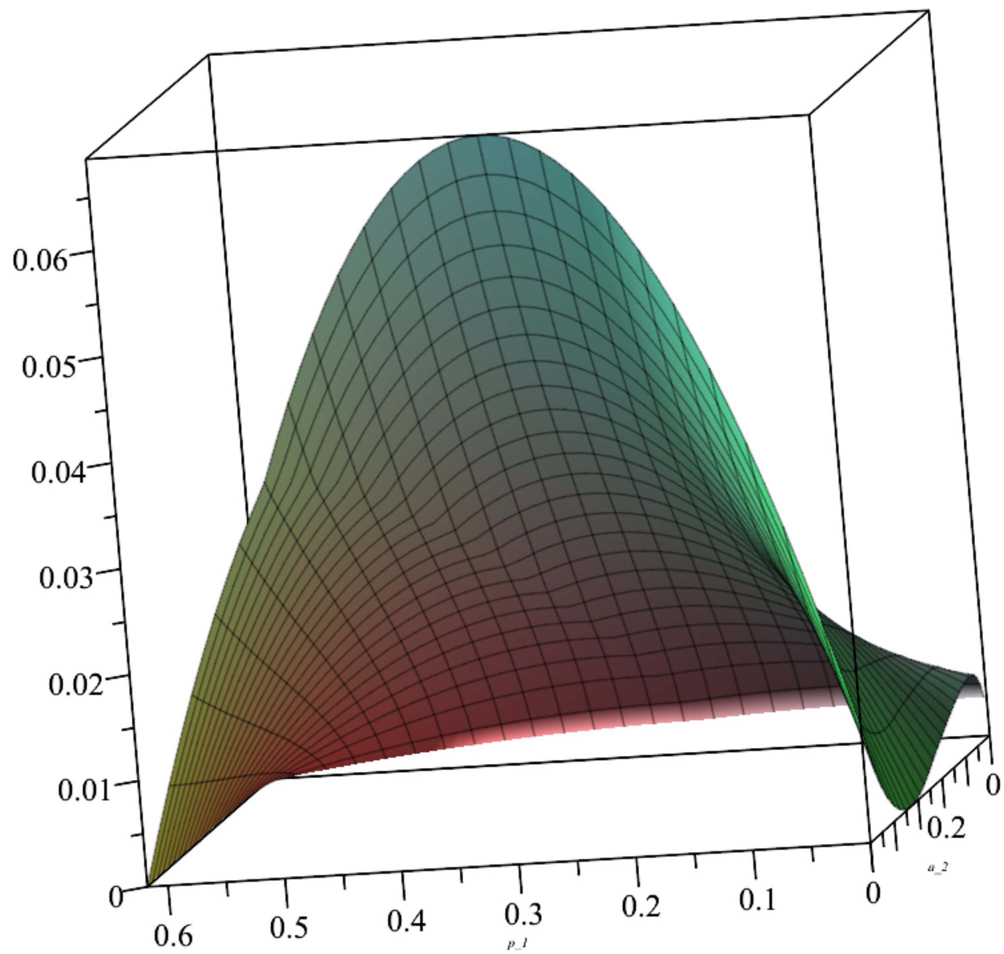
```
>
```

```
> plots:-implicitplot3d(F(p_1, a_1, (p_1·a_1 + 1 - p_1 - C)/(1 - a_2), a_2) - 0.00001 , a_1 = 0 .. 0.25, p_1 = 0 ..1 - C, a_2 = 0 .. 0.5, numpoints = 10000, style = surface, color = navy);
```



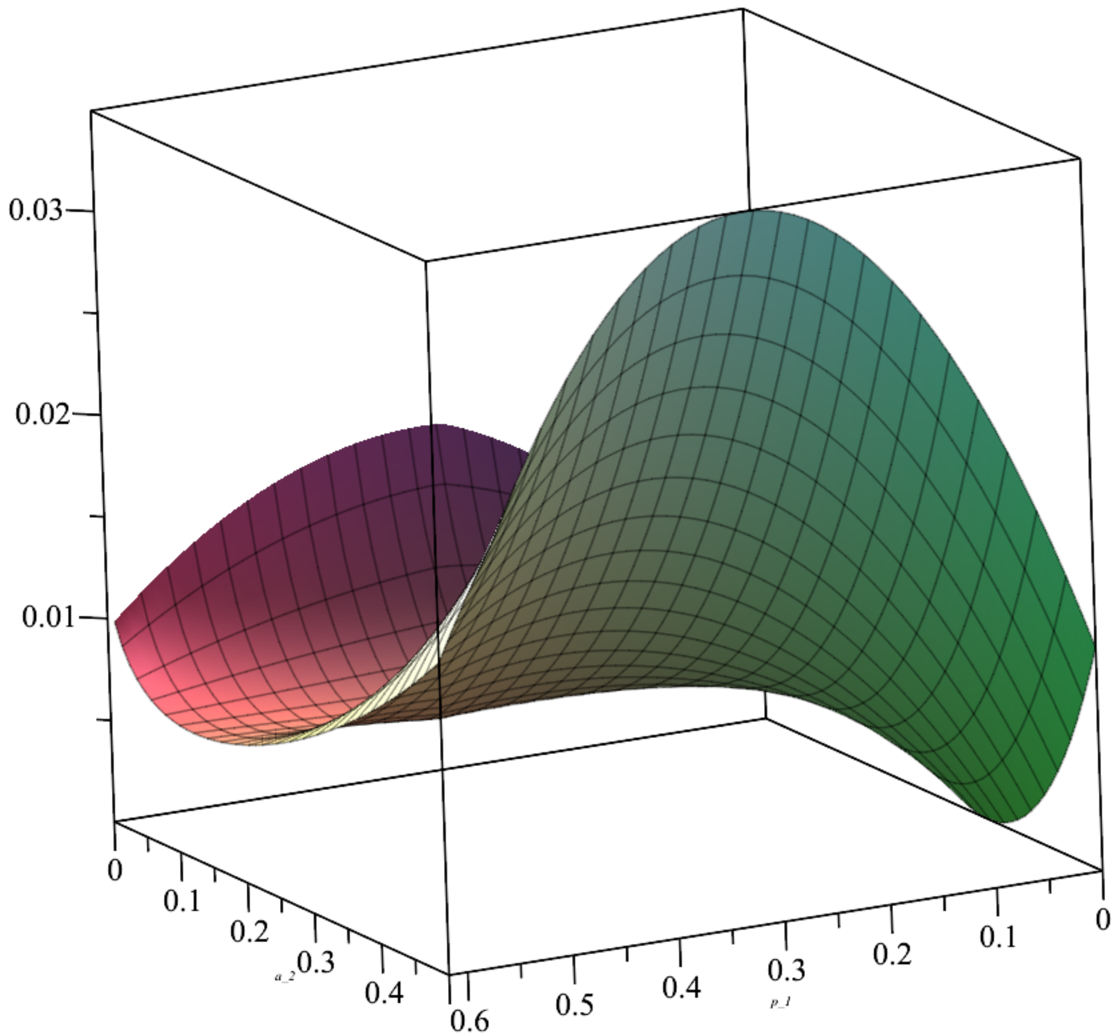
(8)

> $x := 0.00001;$
> $plot3d\left(F\left(p\_1, x, \frac{(1 - p\_1 \cdot (1 - x) - C)}{1 - a\_2}, a\_2\right), p\_1 = 0 ..1 - C, a\_2 = 0 .. 0.5\right)$



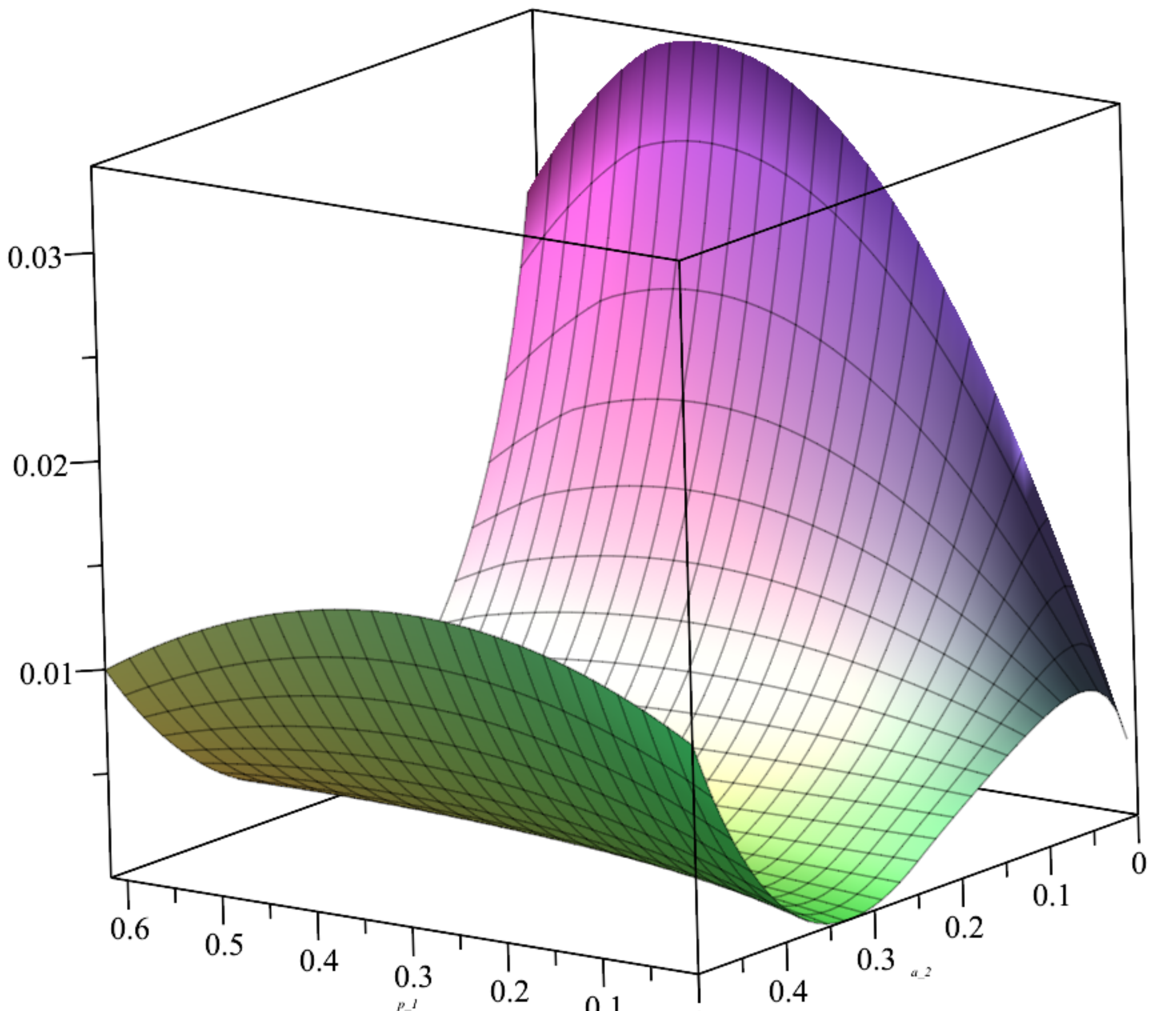
> $x := 0.1;$
$plot3d\left(F\left(p\_1, x, \frac{(1-p\_1\cdot(1-x)-C)}{1-a\_2}, a\_2\right), p\_1=0..1-C, a\_2=0..0.5\right)$

$$x := 0.1$$



> $x := 0.25;$
> $plot3d\left(F\left(p\_1, x, \frac{(1 - p\_1 \cdot (1 - x) - C)}{1 - a\_2}, a\_2\right), p\_1 = 0..1 - C, a\_2 = 0..0.5\right)$

$$x := 0.25$$



0.03
0.02
0.01
0.6
0.5
0.4
0.3
0.2
0.1
0
p_1
0.4
0.3
0.2
0.1
0
a_2