

```

1 local help= []
2 NB:
3 (c)2022 Tim Menzies, timm@ieee.org
4
5 OPTIONS:
6 --k --k handle rare classes = 1
7 --m --m handle rare attributes = 2
8 --p --p distance coefficient = 2
9
10 OPTIONS (other):
11 -h --help show help = false
12 -g --go start-up goal = nothing
13 -s --seed seed = 10019
14 -f --f file = ../../data/auto93.csv]]
15 -----
16 --
17 --
18
19 local lib = require"lib"
20 local cli, csv, demo, is, normpdf = ilib.cli, lib.csc, lib.demo, lib.is, lib.normpdf
21 local o, read, str = lib.o, lib.read, lib.str
22
23 local THE={}
24 help:gsub("[^-][^%s+][^w]%%s[|^%s+]", function(key,x) THE[key]=read(x) end)
25
26 local NUM, SYM, COLS, ROWS = is"NUM", is"SYM", is"COLS", is"ROWS"
27 -----
28 --
29 --
30 --
31 function add(i, x)
32   for _,v in pairs(type(x)=="table" and x or {x}) do
33     if v ~= "" then
34       i.n = i.n + 1
35       i:add(v) end end
36   return x end
37
38 function NUM.new(i) i.n,i.mu,i.m2,i.mu = 0,0,0,0 end
39 function NUM.mid(i,p) return rnd(i.mu,p) end
40 function NUM.like(i,x,...) return normpdf(x, i.mu, i.sd) end
41 function NUM.add(i,v)
42   d = v - i.mu
43   i.mu = i.mu + d/i.n
44   i.m2 = i.m2 + d*(v - i.mu)
45   i.sd = i.n<2 and 0 or (i.m2/(i.n-1))^0.5 end
46
47 function SYM.new(i) i.n,i.syms,i.most,i.mode = 0,{},0,nil end
48 function SYM.mid(i,...) return i.mode end
49 function SYM.like(i,x,prior) return ((i.syms[x] or 0)+THE.m*prior)/(i.n+THE.m) end
50 function SYM.add(i,v)
51   i.syms[v] = (inc or 1) + (i.syms[v] or 0)
52   if i.syms[x] > i.most then i.most,i.mode = i.syms[v],v end end
53 -----
54 --
55 --
56 --
57 local function usep(x) return not x:find"$" end
58 local function nump(x) return x:find"[A-Z]" end
59 local function goalp(x) return x:find"[+-]" end
60 local function klassp(x) return x:find"$" end
61
62 function new(at,txt)
63   txt = txt or ""
64   local i = (nump(txt) and NUM or SYM)()
65   i.txt, i.usep, i.at, i.w = txt, usep(txt), at or 0, txt:find"$" and -1 or 1
66   return i end
67
68 function COLS.new(i,t)
69   i.all, i.xs, i.ys, i.names = {}, {}, {}, t
70   for at,x in pairs(t) do
71     col = push(i.all, new(at,x))
72     if col.usep then
73       if klassp(col.txt) then i.klass=col end
74       push(goalp(col.txt) and i.ys or i.xs, col) end end end
75
76 function COLS.add(i,t)
77   for _,cols in pairs(i.xs,i.ys) do
78     for _,col in pairs(cols) do col:add(t[col.at]) end end
79   return t end
80 -----
81 --
82 --
83 --
84 local function load(src, fun)
85   if type(src)=="string" then for _,t in pairs(src) do fun(t) end
86   else for t in csv(src) do fun(t) end end end
87
88 function ROWS.new(i,t) i.cols=COLS(t); i.rows={} end
89 function ROWS.add(i,t) push(i.rows, i.cols:add(t)) end
90 function ROWS.mid(i, p)
91   t={}; for k,v in pairs(i.cols.ys) do t[k]=col:mid(p) end; return t end
92
93 function ROWS.clone(i,t, j)
94   j= ROWS((i.cols.names)); for _,row in pairs(t) do j:add(row) end; return j end
95
96 function ROWS.like(i,t, nklases, nrows, prior, like, inc, has)
97   prior = (i.n + THE.k) / (nrows + THE.k * nklases)
98   like = math.log(prior)
99   for _,col in pairs(i.cols.xs) do
100     x = t[col.at]
101     if x and x ~= "?" then
102       like = like + math.log(col:like(x,prior)) end end
103   return like end

```

```

104 -----
105 --
106 --
107 --
108 local no, go = {}, {}
109 function go.csv() for row in csv(THE.file) do oo(row) end end
110
111 -----
112 --
113 --
114 --
115 if pcall(debug.getlocal, 4, 1)
116 then return {ROW=ROW, ROWS=ROWS, NUM=NUM, SYM=SYM, THE=THE, lib=lib}
117 else THE = cli(THE, help)
118   demos(THE, go) end

```