

```

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 -b --help show help = false
12 -g --go start-up goal = nothing
13 -s --seed seed = 10019
14 -f --file file = ../../data/auto93.csv]]
15 -----
16 -- 1 0 1 1 1 0 1 5
17
18 local lib = require"lib"
19 local cli, csv, demos, is, normpdf = lib.cli, lib.csv, lib.demos, lib.is, lib.normpdf
20 local oo, read, str = lib.oo, lib.read, lib.str
21
22 local THE={}
23 help:gsub("[^-][^%s+][^%n]*%s{[^%s+]*}", function(key,x) THE[key] = read(x) end)
24
25
26
27 local NUM, SYM, COLS, ROWS = is"NUM", is"SYM", is"COLS", is"ROWS"
28 -----
29 -- 2 0 1 1 1 1 1 5
30
31 local 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 -- 2 0 1 1 1 1 1 5
55
56 local function usep(x) return not x:find"$" end
57 local function nump(x) return x:find"*[A-Z]" end
58 local function goalp(x) return x:find"[+-]$" end
59 local function klassp(x) return x:find"$" end
60
61 local function new(at, txt)
62 txt = txt or ""
63 local i = (nump(txt) and NUM or SYM)()
64 i.txt, i.usep, i.at, i.w = txt, usep(txt), at or 0, txt:find"$" and -1 or 1
65 return i end
66
67 function COLS.new(i, t)
68 i.all, i.xs, i.ys, i.names = {}, {}, {}, t
69 for at, x in pairs(t) do
70 col = push(i.all, new(at, x))
71 if col.usep then
72 if klassp(col.txt) then i.klass=col end
73 push(goalp(col.txt) and i.ys or i.xs, col) end end end
74
75 function COLS.add(i, t)
76 for _, col in pairs(i.xs, i.ys) do
77 col:add(t[col.at]) end end
78 return t end
79 -----
80 -- 1 0 1 1 1 5
81
82 local function load(src, fun)
83 if type(src)=="string" then for _, t in pairs(src) do fun(t) end
84 else for t in csv(src) do fun(t) end end end
85
86 function ROWS.new(i, t) i.cols=COLS(t); i.rows={} end
87 function ROWS.add(i, t) push(i.rows, i.cols:add(t)) end
88 function ROWS.mid(i, p)
89 t={}; for k,v in pairs(i.cols.ys) do t[k]=col:mid(p) end; return t end
90
91 function ROWS.clone(i, t, j)
92 j= ROWS((i.cols.names)); for _, row in pairs(t) do j:add(row) end; return j end
93
94 function ROWS.like(i, t, nklases, nrow, prior, like, inc, has)
95 prior = (i.n + THE.k) / (nrow + THE.k * nklases)
96 like = math.log(prior)
97 for _, col in pairs(i.cols.xs) do
98 x = t[col.at]
99 if x and x ~= "?" then
100 like = like + math.log(col:like(x, prior)) end end
101 return like end
102
103
104
105 -----
106 -- 1 1 0 1 1 5
107
108 local no, go = {}, {}
109 function go.the() oo(THE); return true end
110 function go.csv() print(THE.file); for row in csv(THE.file) do oo(row) end; return tr
111 ue end
112 -----
113 -- 1 1 0 1 1 5
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 oo(THE)
119 demos(THE, go) end
120
121

```