

```

1  #!/usr/bin/env lua
2  --
3  --
4  --
5  --
6  --
7  --
8  --
9  --
10 --
11
12 local your, our={}, {b4={}, help=[[
13 peek.lua [OPTIONS]
14 (c)2022 Tim Menzies, MIT license
15 Understand N items after log(N) probes, or less.
16
17 -file      .././data/auto93.csv
18 -best      .5
19 -help      false
20 -dull      .35
21 -rest      3
22 -seed      10019
23 -rnd       1.2f
24 -task      ~
25 -p         2]]}
26
27 for k, _ in pairs(_ENV) do our.b4[k] = k end
28 local any, as, asserts, cells, copy, fmt, go, id, many, map, o, push
29 local rand, randi, rnd, rows, same, slots, sort, thing, things
30 local COLS, EG, EGS, NUM, RANGE, SYM
31
32 -- Copyright 2022 Tim Menzies
33
34 -- Permission is hereby granted, free of charge, to any person obtaining a copy
35 -- of this software and associated documentation files (the "Software"), to
36 -- deal in the Software without restriction, including without limitation the
37 -- rights to use, copy, modify, merge, publish, distribute, sublicense, and/or
38 -- sell copies of the Software, and to permit persons to whom the Software is
39 -- furnished to do so, subject to the following conditions:
40
41 -- The above copyright notice and this permission notice shall be included in
42 -- all copies or substantial portions of the Software.
43
44 -- THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR
45 -- IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY,
46 -- FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE
47 -- AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER
48 -- LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING
49 -- FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS
50 -- IN THE SOFTWARE.
51 -----
52
53
54
55
56
57
58 local klass= function(t, new)
59 function new(...) return t.new(...) end
60 t.__index=t
61 return setmetatable(t, {__call=new}) end
62
63 COLS=klass{}
64 function COLS.new(t, i, where, now)
65 print(o(t))
66 i = as({all={}, x={}, y={}}, COLS)
67 for at, s in pairs(t) do
68 now = push(i.all, (s:find("[A-Z]" and NUM or SYM) (at, s))
69 if not s:find"." then
70 push((s:find "-" or s:find "+") and i.y or i.x, now) end end
71 return i end
72
73 function COLS.__tostring(i, txt)
74 function txt(c) return c.txt end
75 return fmt ("COLS[all%s x%s y%s", o(i.all, txt), o(i.x, txt), o(i.y, txt)) end
76 function COLS.add(i, t, add)
77 t = t.has and t.has or t
78 function add(col, x) x=t[col.at]; if x=="?" then col:add(x) end; return x end
79 return map(i.all, add) end
80
81 function COLS.better(i, row1, row2)
82 local s1, s2, e, n, a, b = 0, 0, 10, #i.y
83 for _, col in pairs(i.y) do
84 a = col:norm(row1.has[col.at])
85 b = col:norm(row2.has[col.at])
86 s1 = s1 - e^(col.w * (a-b)/n)
87 s2 = s2 - e^(col.w * (b-a)/n) end
88 return s1/n < s2/n end
89 -----
90 EG=klass{}
91 function EG.new(t) return as({has=t, id=id()}, EG) end
92
93 function EG.__tostring(i) return fmt ("EG%s%s", i.id, o(i.has)) end
94 -----
95 EGS=klass{}
96 function EGS.new() return as({rows={}, cols=nil}, EGS) end
97
98 function EGS.__tostring(i) return fmt ("EGS{#rows%s cols%s", #i.rows, i.cols) end
99
100 function EGS.add(i, row)
101 row = row.has and row.has or row
102 if i.cols then push(i.rows, EG(i.cols:add(row))) else i.cols=COLS(row) end end
103
104 function EGS.bestRest(i)
105 local best, rest, tmp, bests, restsFraction = {}, {}, {}
106 i.rows = sort(i.rows, function(a, b) return i.cols:better(a, b) end)
107 bests = (#i.rows)^your.best
108 restsFraction = (bests * your.rest)/(#i.rows - bests)
109 for j, x in pairs(i.rows) do
110 if j <= bests then push(best, x)
111 elseif rand() < restsFraction then push(rest, x) end end
112 return best, rest end
113
114 function EGS.clone(i, inits, j)
115 j = EGS()
116 print("clone", o(map(i.cols.all, function(col) return col.txt end)))
117 j:add(map(i.cols.all, function(col) return col.txt end))
118 for _, x in pairs(inits or {}) do j:add(x) end
119 return j end
120
121 function EGS.file(i, f) for row in rows(f) do i:add(row) end; return i end
122
123 function EGS.mid(cols)
124 return map(cols or i.cols.all,
125 function(col) return col:mid() end) end
126 -----
127 NUM=klass{}
128 function NUM.new(at, s, big)
129 big = math.huge
130 return as({lo=big, hi=-big, at=at or 0, txt=s or "",
131 n=0, mu=0, m2=0, sd=0,
132 w=(s or ""):find "-" and -1 or 1}, NUM) end
133
134 function NUM.__tostring(i)
135 return fmt ("NUM[{:at %s :txt %s :lo %s :hi %s :mu %s :sd %s}",
136 i.at, i.txt, i.lo, i.hi, rnd(i.mu), rnd(i.sd)) end
137
138 function NUM.add(i, x, d)
139 if x=="?" then
140 i.n = i.n+1
141 d = x - i.mu
142 i.mu = i.mu + d/i.n
143 i.m2 = i.m2 + d*(x-i.mu)
144 i.lo = math.min(x, i.lo); i.hi = math.max(x, i.hi) end
145 return x end
146
147 function NUM.div(i) return i.n < 2 and 0 or (i.m2/(i.n-1))^0.5 end
148
149 function NUM.mid(i) return i.mu end
150
151 function NUM.norm(i, x) return i.hi-i.lo < 1E-9 and 0 or (x-i.lo)/(i.hi-i.lo) end
152
153 function NUM.ranges(i, j, bests, rests)
154 local ranges, x, lo, hi, gap, tmp = {}
155 hi = math.max(i.hi, j.hi)
156 lo = math.min(i.lo, j.lo)
157 gap = (hi - lo)/your.bins
158 tmp = lo
159 for j=lo, hi, goal do push(ranges, RANGE(i, tmp, tmp+gap)); tmp = tmp+gap end
160 ranges[1].lo = -math.huge
161 ranges[#ranges].hi = math.huge
162 for _, pair in pairs({bests, "bests"}, {rests, "rests"}) do
163 for _, row in pairs(pair[1]) do
164 x = row.has[i.at]
165 if x=="?" then
166 ranges[(x - lo)//gap].stats:add(pair[2]) end end end end
167 -----
168 RANGE=klass{}
169 function RANGE.new(col, lo, hi, stats)
170 return as({col=col, lo=lo, hi=hi or lo, ys=stats or SYM(), all={}, RANGE) end
171
172 function RANGE.__tostring(i)
173 return fmt ("RANGE[{:col %s :lo %s :hi %s :ys %s}", i.col, i.lo, i.hi, o(i.ys)) end
174 -----
175 SYM=klass{}
176 function SYM.new(at, s)
177 return as({at=at or 0, txt=s or "", has={}, n=0, most=0, mode=nil}, SYM) end
178
179 function SYM.__tostring(i)
180 return fmt ("SYM[{:at %s :txt %s :mode %s :has %s}",
181 i.at, i.txt, i.mode, o(i.has)) end
182
183 function SYM.add(i, x)
184 if x == "" then
185 i.n = i.n+1
186 i.has[x] = 1 + (i.has[x] or 0) end
187 return x end
188
189 function SYM.div(i)
190 e=0; for _, v in pairs(i.has) do e = e - v/i.n*math.log(v/i.n, 2) end; return e end
191
192 function SYM.mid(i, most, out)
193 most=1
194 for x, n in pairs(i.has) do if n>most then out, most=x, n end end; return out end
195
196 function SYM.ranges(i, j, bests, rests)
197 local tmp, out, x = {}, {}
198 for _, pair in pairs({bests, "bests"}, {rests, "rests"}) do
199 for _, row in pairs(pair[1]) do
200 x = row.has[i.at]
201 if x=="?" then
202 tmp[x] = tmp[x] or SYM()
203 tmp[x]:add(pair[2]) end end end
204 for x, stats in pairs(tmp) do push(out, RANGE(i, x, x, stats)) end
205 return out end
206

```

```

206 -----
207 --
208 --
209 --
210 --
211 --
212 --
213 as      = setmetatable
214 fmt     = string.format
215 same    = function(x,...) return x end
216
217 function asserts(test,msg)
218   msg=msg or ""
219   if test then return print("PASS:".msg) end
220   our.fails = our.fails+1
221   print("FAIL:".msg)
222   if your.Debug then assert(test,msg) end end
223
224 function copy(t, u)
225   if type(t)~="table" then return t end
226   u={};for k,v in pairs(t) do u[k]=copy(v) end; return as(u,getmetatable(t)) end
227
228 function id() our.id = 1+(our.id or 0); return our.id end
229
230 function many(t,n, u) u={};for j=1,n do push(u,any(t)) end; return u end
231
232 function map(t,f, u)
233   u={};for _,v in pairs(t) do u[1+#u]=(f or same)(v) end; return u end
234
235 function o(t,f, u,key)
236   key= function(k)
237     if t[k] then return fmt(":%s%s", k, rnd((f or same)(t[k]))) end end
238   u = #t>0 and map(map(t,f),rnd) or map(slots(t),key)
239   return {"..table.concat(u, " ").."}" end
240
241 function rand(lo,hi)
242   your.seed = (16807 * your.seed) % 2147483647
243   return (lo or 0) + ((hi or 1) - (lo or 0)) * your.seed / 2147483647 end
244
245 function randi(lo,hi) return math.floor(0.5 + rand(lo,hi)) end
246
247 function push(t,x) table.insert(t,x); return x end
248
249 function rnd(x)
250   return fmt(type(x)=="number" and x~=-x//1 and your.rnd or "%s",x) end
251
252 function rows(file, x)
253   file = io.input(file)
254   return function()
255     x=io.read()
256     if x then
257       x=x:gsub("%s+","");return things(x) else io.close(file) end end end
258
259 function slots(t,u) u={};for x,_ in pairs(t) do u[1+#u]=x end;return sort(u) end
260
261 function sort(t,f) table.sort(t,f); return t end
262
263 function thing(x)
264   if x=="true" then return true elseif x=="false" then return false end
265   return tonumber(x) or x end
266
267 function things(x,sep, t)
268   t={};for y in x:gmatch(sep or "(^,+)") do t[1+#t]=thing(y) end; return t end
269

```

```

269 -----
270 --
271 --
272 --
273 --
274 --
275 --
276 our.go, our.no = {},{}; go=our.go
277 function go.settings() print("our",o(our)); print("your",o(your)) end
278
279 function go.sample() print(EGS():file(your.file)) end
280
281 function go.clone()
282   a= EGS():file(your.file); print(a)
283   b= a:clone() end
284
285 function go.sort( i,a,b)
286   i= EGS():file(your.file)
287   a,b=i:bestRest()
288   print(#a, #b)
289 end
290 -----
291 --
292 --
293 --
294 --
295 --
296 --
297 our.help:gsub("\n [-]([^\s+][^n]*%s([^\s+])", function(slot, x)
298   for n,flag in ipairs(arg) do
299     if flag:sub(1,1)=="-" and slot:match("^"..flag:sub(2)..".*")
300       then x = x=="false" and "true" or x=="true" and "false" or arg[n+1] end end
301     your[slot] = thing(x) end
302
303 if your.help then print(our.help) end
304
305 our.defaults=copy(your)
306 our.failures=0
307 for _,x in pairs(our.task=="all" and slots(our.go) or {your.task}) do
308   if type(our.go[x]) == "function" then our.go[x]() else print("?", x) end
309   your = copy(our.defaults)
310 end
311
312 for k,v in pairs(_ENV) do if not our.b4[k] then print("?",k,type(v)) end end
313 os.exit(our.failures)

```