

```

1  b4={}; for k,_ in pairs(_ENV) do b4[k]=k end
2  local r,abs,log,ent,min,max
3  local copy,push,fmt,fmt2,map,map2,cat,cat2,rnd,rnds
4  local add,cls,thing,things,cls
5  local ok,cli,demos,demo
6  local the,fails,go,no
7  local Num,Sym,Cols,Egs
8
9  the={k      = 2,
10       m      = 1,
11       seed    = 10019,
12       rnd     = "%5.3f",
13       dump    = false,
14       todo    = "the",
15       keep    = 512}
16
17 -----
18 r= math.random
19 abs= math.abs
20 log= math.log
21 min= math.min
22 max= math.max
23 push= function(t,x) t[1 + #t] = x; return x end
24
25 fmt= string.format
26 fmt2= function(k,v) return fmt(":%s %s",k,v) end
27
28 map= function(t,f,u) u={};for _,v in pairs(t)do u[1+#u]=f(v) end;return u end
29 map2= function(t,f,u) u={};for k,v in pairs(t)do u[1+#u]=f(k,v)end;return u end
30
31 copy= function(t, u)
32   if type(t) ~= "table" then return t end
33   u={};for k,v in pairs(t) do u[copy(k)]=copy(v) end; return u end
34
35 cat= function(t) return {"..table.concat(map(t,tostring), " ").".."} end
36 cat2= function(t,sep) return {"..table.concat(map2(t,fmt2),sep or " ").".."} end
37
38 rnd= function(x,f)
39   return fmt(type(x)=="number" and (x~x//1 and f or the.rnd) or "%s",x) end
40 rnds= function(t,f) return map(t, function(x) return rnd(x,f) end) end
41
42 ent= function(t, n,e)
43   n=0; for _,v in pairs(t) do n=n+v end
44   e=0; for k,v in pairs(t) do e=e-v/n*log(v/n,2) end; return e end
45
46 thing= function(x)
47   x = x:match"%s*(-)%s*$"
48   if x=="true" then return true elseif x=="false" then return false end
49   return math.tointeger(x) or tonumber(x) or x end
50
51 things= function(s,sep, t)
52   t={}; for y in s:gmatch("([^\,]+)") do t[1+#t]=coerce(y) end
53   return t end
54
55 csv= function(src)
56   src = io.input(src)
57   return function(x) x=io.read()
58     if x then return things(x) else io.close(src) end end end
59
60 class= function(name, t,new)
61   function new(klass,...)
62     local obj= setmetatable({},klass)
63     local res= klass.new(obj,...)
64     if res then obj = setmetatable(res,klass) end
65     return obj end
66   t[("__tostring",cat2, _is=name or ""); t.__index=t
67     return setmetatable(t, {__call=new}) end
68
69 add= function(obj,data)
70   if type(data)=="string"
71   then for row in csv(data) do obj:add(row) end
72   else for _,row in pairs(data or {}) do obj:add(row) end end
73   return obj end
74
75 cli= function(the, k,v)
76   for n,flag in ipairs(arg) do
77     k = flag:sub(3)
78     v = the[k]
79     if v ~= nil then
80       v = v==false and "true" or v==true and "false" or arg[n+1]
81       the[k] = thing(v) end end
82   return the end
83
84 fails=0
85 ok= function(test,msg)
86   print("", test and "PASS" or "FAIL",msg or "")
87   if not test then
88     fails= fails+1
89     if the.dump then assert(test,msg) end end end
90
91 demo= function(txt,f,old)
92   the = copy(old)
93   math.randomseed(the.seed or 10019)
94   print(txt)
95   f() end
96
97 demos= function(the,go, old)
98   old = copy(the)
99   if the.todo=="all"
100  then for s,f in pairs(go) do demo(s, f, old) end
101  else demo(the.todo, go[the.todo], old) end end
102
103 -----
104 Num=class("Num")
105 function Num:new(at,name)
106   self.at, self.name = at or 0, name or ""
107   self.w = self.name:find"$-" and -1 or 1
108   self.some, self.ok = {}, false
109   self.n,self.md,self.sd,self.lo,self.hi = 0,0,0,1E32,-1E32 end
110
111 function Num:add(x,_, a,d)
112   if x ~="?" then
113     self.n = self.n + 1
114     d = x - self.mu
115     self.mu= self.mu + d/self.n
116     self.m2= self.m2 + d*(x - self.mu)
117     self.sd= (self.m2<0 or self.n<2) and 0 or ((self.m2/(self.n - 1))^0.5)
118     self.lo= min(x, self.lo)
119     self.hi= max(x, self.hi)
120     a = self.some
121     if #a < the.num.keep then self.ok=false; push(a,x)
122     elseif r() < the.num.keep/self.n then self.ok=false; a[r(#a)]=x end end
123   return x end
124
125 function Num:mid() return self.mu end
126 function Num:div() return self.sd end
127
128 function Num:like(x,_)
129   local z, e, pi = 1E-64, math.exp(1), math.pi
130   if x < self.mu - 4*self.sd then return 0 end
131   if x > self.mu + 4*self.sd then return 0 end
132   return e^(-(x - self.mu)^2 / (z + 2*self.sd^2))/(z + (pi*2*self.sd^2)^.5) end
133
134 -----
135 Sym=class("Sym")
136 function Sym:new(at,name)
137   self.at, self.name = at or 0, name or ""
138   self.has, self.mode, self.most = {},nil,0 end
139
140 function Sym:add(x,inc)
141   if x ~="" then
142     inc = inc or 1
143     self.n = self.n + inc
144     self.has[x] = inc + (self.has[x] or 0)
145     if self.has[x] > self.most then
146       self.most, self.mode = self.has[x], x end end
147   return x end
148
149 function Sym:mid() return self.mode end
150 function Sym:div() return ent(self.has) end
151
152 function Sym:like(x,prior)
153   return ((self.has[x] or 0) + the.m*prior)/(self.n + the.m) end
154
155 -----
156 Cols=class("Cols")
157 function Cols:new(names, col)
158   self.names = names
159   self.all, self.x, self.y = {}, {}, {}
160   for at,name in pairs(names) do
161     col = push(self.all, (name:find"^[A-Z]" and Num or Sym) (at,name))
162     if not name:find"$" then
163       if name:find"$" then self.klass=col end
164       col.indep = not name:find"[+!]"
165       push(col.indep and self.x or self.y, col) end end end
166
167 -----
168 Egs=class("Egs")
169 function Egs:new() self.rows, self.cols = {},nil end
170
171 function Egs:add(row, add)
172   add = function(col) col:add(row[col.at]) end
173   if self.cols then push(self.rows, map(self.cols,add)) else
174     self.cols = Cols(row) end end
175
176 function Egs:mid(cols)
177   return map(cols or self.cols.y, function(col) return col:mid() end) end
178
179 function Egs:div(cols)
180   return map(cols or self.cols.y, function(col) return col:mid() end) end
181
182 function Egs:like(row,egs, n,prior,like,col)
183   n=0; for _,eg in pairs(egs) do n = n + #eg.rows end
184   prior = (#self.rows + the.k) / (n + the.k * #egs)
185   like = log(prior)
186   for at,x in pairs(row) do
187     col = self.cols.all[at]
188     if x ~="?" and col.indep then like= like + log(col:like(x,prior)) end end
189   return like end
190
191 -----

```

```
191 -----
192 go,no = {}, {}
193 function go.the() print(cat2(the)) end
194
195 the = cli(the)
196 demos(the,go)
197
198 for k,v in pairs(_ENV) do if not b4[k] then print("?",k,type(v)) end end
```