

```

1 local _, the, COL = require"lib", require"the", require"col"
2 local class,merge,per,push,sort,upx = _.class,_.merge,_.per,_.push,_.sort,_.upx
3 local sd = _.sd
4 local norm,oo = _.norm,_.oo
5
6 local NUM = class("NUM",COL)
7 function NUM:new(at,name)
8   self:super(at,name)
9   self.has, self.ok = {}, false
10  self.lo, self.hi = math.huge, -math.huge end
11
12 local r=math.random
13 function NUM:add1(x,inc, pos)
14   for i=1,inc do
15     self.lo = math.min(x, self.lo)
16     self.hi = math.max(x, self.hi)
17     if #self.has < the.some then pos = 1 + #self.has
18     elseif r() < the.some/self.n then pos = 1 + ((r()*#self.has)//1) end
19     if pos then
20       self.ok = false
21       self.has[pos] = x end end end
22
23 function NUM:div(a) a=self:all(); return (per(a,.9) - per(a,.1))/2.56 end
24 function NUM:mid() return per(self:all(), .5) end
25 function NUM:same(x,y) return math.abs(x - y) <= the.cohen * self:div() end
26
27 function NUM:dist1(x,y)
28   if x=="?" then y = norm(self.lo, self.hi, y); x=y<.5 and 1 or 0
29   elseif y=="?" then x = norm(self.lo, self.hi, x); y=x<.5 and 1 or 0
30   else
31     x,y = norm(self.lo, self.hi, x), norm(self.lo, self.hi,y) end
32   return math.abs(x-y) end
33
34 function NUM:like1(i,x)
35   local sd= self:div()
36   if x < self.mu - 4*sd then return 0 end
37   if x > self.mu + 4*sd then return 0 end
38   local denom = (math.pi*2*sd^2)^.5
39   local nom = math.exp(1)^(-(x-self.mu)^2/(2*sd^2+1E-32))
40   return nom/(denom + 1E-32) end
41
42 function NUM:merge(other, out)
43   out = NUM(self.at, self.name)
44   for _,x in self(self.has) do out:add(x) end
45   for _,x in self(other.has) do out:add(x) end
46   return out end
47
48 function NUM:all()
49   if not self.ok then table.sort(self.has) end
50   self.ok=true
51   return self.has end
52
53 function NUM:bins(other, BIN)
54   local tmp,out = {},{}
55   for _,x in pairs(self.has) do push(tmp, {x=x, y="left"}) end
56   for _,x in pairs(other.has) do push(tmp, {x=x, y="right"}) end
57   tmp = sort(tmp,upx) -- ascending on x
58   local now = push(out, BIN(self.at, self.name, tmp[1].x))
59   local epsilon = sd(tmp,function(z) return z.x end) * the.cohen
60   local minSize = (#tmp)^the.leaves
61   for j,xy in pairs(tmp) do
62     if j > minSize and j + minSize < #tmp then -- leave enough for other bins
63       if now.ys.n > minSize then -- enough in this bins
64         if xy.x ~- tmp[j+1].x then -- there is a break in the data
65           if now.hi - now.lo > epsilon then -- "now" not trivially small
66             now = push(out, BIN(self.at, self.name, now.hi)) end end end end
67         out[1].lo = -math.huge
68         out[#out].hi = math.huge
69         return merge(out, BIN.mergeSameDivs) end
70   end
71   return NUM

```