Subject: Calculating cumulative probability using cgHistoPlot

Posted by Xin Tao on Tue, 08 Nov 2011 18:10:03 GMT

View Forum Message <> Reply to Message

Hi all,

I'm using cghistoplot, and I'm confused by the cumulative probability calculated by cghistoplot. I'm wondering whether any one can give me some help here.

Suppose we have histdata like this: [2, 0 , 0, 3, 5, 4], then if we use the way of cghistoplot to calculate the cumulative probability like this:

```
  cumTotal = Total(histData, /CUMULATIVE)   ;; gives us
[2.00000,    2.00000,    2.00000,    5.00000,    10.0000,
14.0000]
  probability = Scale_Vector(cumTotal, 0, 1)    ;; gives us
probability = [0.00000     0.00000     0.00000    0.250000
0.666667     1.00000]
```

This is kind of counter-intuitive to me, because the first value of histdata is clearly 2, but the probability is 0 until the 4th value. However, I'm not experienced in data anlaysis, and I might have misunderstood something about "cumulative probability" here.

It seems to be more natural to me to define the cumulative probability in the following way:

probability = total(double(hist)/total(double(hist)), /cumula)

Am I right?

Xin Tao