Subject: Re: Calculating cumulative probability using cgHistoPlot
Posted by David Fanning on Tue, 08 Nov 2011 19:57:42 GMT
View Forum Message <> Reply to Message

Xin Tao writes:

>
> 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?

I think you are right. I was both calculating this incorrectly
and displaying it incorrectly. I think you will be more pleased
with the updated program. :-)

http://www.idlcoyote.com/programs/cghistoplot.pro

Thanks for pointing this error out.

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.idlcoyote.com/
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

---