Subject: Re: significance probability when using CORRELATE?
Posted by mchinand on Wed, 19 Sep 2001 19:43:03 GMT
View Forum Message <> Reply to Message

In article <3BA8EBD3.DF2D3C6A@hpl.umces.edu>,
Linda May <may@hpl.umces.edu> wrote:
> I work with oceanographic data using IDL 5.4 (on a Compaq Alpha running
> Tru64
> Unix). Often, in preliminary analyses, I use CORRELATE between two sets
> of
> data. For example, is there a linear association between wind and sea
> surface
> temperature? However, it is important to know not only the correlation
> coefficient, but also its significance probability.
>
> In SAS, when I ask for the (linear Pearson) correlation coefficient of
> two
> vectors, I am automatically given the significance probability of the
> correlation. The (two-tailed) probability value reported by SAS comes
> from
> a t distribution (symmetric with respect to the y-axis).
>
> But I haven't been able to figure out how to obtain the significance
> probability
> in IDL. I've investigated such routines as TM_TEST, but I am not a
> statistician
> and it's not clear to me how I could use TM_TEST. The significance
> probability
> seems so important to anyone using CORRELATE that surely there must be
> an
> obvious and simple way to obtain it in IDL which I am overlooking. I've
>
> searched a number of libraries, but haven't found any examples.
>
> So I sheepishly tell you that for two years I've been switching back and
>
> forth between IDL and SAS. This has become unbearably tacky in my view;
> thus, I
> would be grateful if you could steer me in a helpful direction.
>
> Cheers,
> Linda
>

For the calculation of the two-tailed significance (Ho: r=0) of a
correlation coefficient I have been using the following code snippet:

r=correlate(x,y)

```
var=1/(n-3.0)
zvalue = 0.5*alog((1+r)/(1-r))/sqrt(var)
if (zvalue lt 0) then pvalue = 2*(gauss_pdf(zvalue)) $
else pvalue = 2*(1-gauss_pdf(zvalue))
```

You should verify that this is the value you want by comparing this value
with your value from SAS. You mentioned a t-distribution in your post, so
I'm not sure if you want something different or not.

Hope that helps,

--Mike

--
--
Michael Chinander   University of Chicago
m-chinander@uchicago.edu   (773)834-5101 (Voice)   (773)702-0371 (Fax)

_____

_____

Subject: Re: significance probability when using CORRELATE?
Posted by mchinand on Wed, 19 Sep 2001 19:48:09 GMT
View Forum Message <> Reply to Message
_____

In article <bz6q7.28$N4.2123@news.uchicago.edu>,
Mike Chinander <mchinand@midway.uchicago.edu> wrote:

> For the calculation of the two-tailed significance (Ho: r=0) of a
> correlation coefficient I have been using the following code snippet:
>
> r=correlate(x,y)
> var=1/(n-3.0)
> zvalue = 0.5*alog((1+r)/(1-r))/sqrt(var)
> if (zvalue lt 0) then pvalue = 2*(gauss_pdf(zvalue)) $
> else pvalue = 2*(1-gauss_pdf(zvalue))
>

I forgot to add, n is the number of x-y pairs.

--
--
Michael Chinander   University of Chicago
m-chinander@uchicago.edu   (773)834-5101 (Voice)   (773)702-0371 (Fax)

_____