Subject: 2D Pearson correlation coefficient
Posted by limiqt on Thu, 30 Jan 2014 14:43:54 GMT
View Forum Message <> Reply to Message

Dear all,
I would like to ask if someone know a code to calculate a 2D Pearson correlation as:

r^2=(Sum wi*(Mi-M)*(Oi-O))^2 /((Sum wi*(Mi-M)^2)*(Sum wi*(Oi-O)^2))

Sum runs from i=1 to N. N is the total number of grid cells.
Mi and Oi are the values in the grid cell i and wi is a normalized weight (area) of grid cell i. Sum wi=1 (Sum from i=1 to N).

IDL has C_Correlate and R_correlate but none of them include the wi factor.

I will appreciate any assistance.

Lim

Subject: Re: 2D Pearson correlation coefficient
Posted by Matthew Argall on Thu, 30 Jan 2014 15:08:45 GMT
View Forum Message <> Reply to Message

> I would like to ask if someone know a code to calculate a 2D Pearson correlation as:
>
> r^2=(Sum wi*(Mi-M)*(Oi-O))^2 /((Sum wi*(Mi-M)^2)*(Sum wi*(Oi-O)^2))
>
> IDL has C_Correlate and R_correlate but none of them include the wi factor.

As far as I know, IDL does not do weighted correlations. It would not take you long to compute it, though. Basically, all you would have to do is replace "Sum" with the Total() function in the above equation...

Subject: Re: 2D Pearson correlation coefficient
Posted by limiqt on Thu, 30 Jan 2014 18:12:13 GMT
View Forum Message <> Reply to Message

I wonder if someone have a suggestion about how to determine the wi (weighting) factor.
Thanks,
Lim

On Thursday, January 30, 2014 9:43:54 AM UTC-5, Lim wrote:
> Dear all,

>
> I would like to ask if someone know a code to calculate a 2D Pearson correlation as:
>
>
>
> r^2=(Sum wi*(Mi-M)*(Oi-O))^2 /((Sum wi*(Mi-M)^2)*(Sum wi*(Oi-O)^2))
>
>
>
> Sum runs from i=1 to N. N is the total number of grid cells.
>
> Mi and Oi are the values in the grid cell i and wi is a normalized weight (area) of grid cell i. Sum wi=1 (Sum from i=1 to N).
>
>
>
> IDL has C_Correlate and R_correlate but none of them include the wi factor.
>
>
>
> I will appreciate any assistance.
>
>
>
> Lim

Subject: Re: 2D Pearson correlation coefficient
Posted by David Fanning on Thu, 30 Jan 2014 18:20:26 GMT
View Forum Message <> Reply to Message

Lim writes:

> I wonder if someone have a suggestion about how to determine the wi (weighting) factor.

Random numbers are always a solid bet, unless you have reason to think
something else would work better.

Cheers,

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

Subject: Re: 2D Pearson correlation coefficient
Posted by Craig Markwardt on Fri, 31 Jan 2014 01:35:42 GMT
View Forum Message <> Reply to Message

On Thursday, January 30, 2014 1:12:13 PM UTC-5, Lim wrote:
> I wonder if someone have a suggestion about how to determine the wi (weighting) factor.

If you don't know, then 1 is the only safe weighting factor.

CM

Subject: Re: 2D Pearson correlation coefficient
Posted by David Fanning on Fri, 31 Jan 2014 01:46:13 GMT
View Forum Message <> Reply to Message

Craig Markwardt writes:

>> I wonder if someone have a suggestion about how to determine the wi (weighting) factor.
>
> If you don't know, then 1 is the only safe weighting factor.

Really!? My first, completely off the cuff answer (unvoiced,
fortunately) was "Geez, just throw the I Ching!" But, I wonder, quite
seriously now, whether a normalized distribution of random numbers isn't
functionally equivalent to using 1. I don't know the answer. But, I
presume you do. :-)

Cheers,

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

Subject: Re: 2D Pearson correlation coefficient
Posted by David Fanning on Fri, 31 Jan 2014 01:52:56 GMT
View Forum Message <> Reply to Message

David Fanning writes:

> Craig Markwardt writes:
>
>>> I wonder if someone have a suggestion about how to determine the wi (weighting) factor.

>>
>> If you don't know, then 1 is the only safe weighting factor.
>
> Really!? My first, completely off the cuff answer (unvoiced,
> fortunately) was "Geez, just throw the I Ching!"

It appears my first answer was closer to the mark than I imagined:

"Because the purpose of Clarity is to make the I Ching's
help available to anyone who needs it, there is absolutely
no need to know or study the I Ching in order to understand
the interpretations I provide.

Whatever you choose, you are invited to make the I Ching's
answers a part of your life - to drink the water from the well."

There you go. :-)

Cheers,

David

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

Subject: Re: 2D Pearson correlation coefficient
Posted by Russell Ryan on Fri, 31 Jan 2014 05:14:40 GMT
View Forum Message <> Reply to Message

Lim,

You're getting snarky responses because you're asking a stupid question. I can understand if you don't know how to calculate the 2d Pearson coefficient. But, how can you expect anyone to have any clue what the weighting coefficients ought to be, since you've told us nothing about what {M} or {O} are? I assume they're data of some sort, but what data? Are they measurements, do they have uncertainties? If so, then what is your error distribution (I mean are they Gaussian uncertainties or Poisson or what). If so, then I'd consider inverse variance weighting, but that's just a hunch.

How can you expect anyone to know what you're doing if you don't tell them?

You should read a few blogs (including David's) on "how to ask a help question." I truly mean no disrespect.

Russell

On Thursday, January 30, 2014 9:43:54 AM UTC-5, Lim wrote:
> Dear all,
>
> I would like to ask if someone know a code to calculate a 2D Pearson correlation as:
>
>
>
> r^2=(Sum wi*(Mi-M)*(Oi-O))^2 /((Sum wi*(Mi-M)^2)*(Sum wi*(Oi-O)^2))
>
>
>
> Sum runs from i=1 to N. N is the total number of grid cells.
>
> Mi and Oi are the values in the grid cell i and wi is a normalized weight (area) of grid cell i. Sum wi=1 (Sum from i=1 to N).
>
>
>
> IDL has C_Correlate and R_correlate but none of them include the wi factor.
>
>
>
> I will appreciate any assistance.
>
>
>
> Lim

Subject: Re: 2D Pearson correlation coefficient
Posted by limiqt on Fri, 31 Jan 2014 09:14:45 GMT
View Forum Message <> Reply to Message

I forgot to mention that M and O are global data of temperature regrided at the same resolution.
Lim

On Thursday, January 30, 2014 9:43:54 AM UTC-5, Lim wrote:
> Dear all,
>
> I would like to ask if someone know a code to calculate a 2D Pearson correlation as:
>
>
>
>

> r^2=(Sum wi*(Mi-M)*(Oi-O))^2 /((Sum wi*(Mi-M)^2)*(Sum wi*(Oi-O)^2))
>
>
>
> Sum runs from i=1 to N. N is the total number of grid cells.
>
> Mi and Oi are the values in the grid cell i and wi is a normalized weight (area) of grid cell i. Sum wi=1 (Sum from i=1 to N).
>
>
>
> IDL has C_Correlate and R_correlate but none of them include the wi factor.
>
>
>
> I will appreciate any assistance.
>
>
>
> Lim

Subject: Re: 2D Pearson correlation coefficient
Posted by robinson.inj on Fri, 31 Jan 2014 11:09:19 GMT
View Forum Message <> Reply to Message

http://www.idlcoyote.com/code_tips/eof_analysis.html

http://stackoverflow.com/questions/10633773/how-to-do-area-weighted-regridding-of-regular-lat-lon-data-using-python

On Thursday, January 30, 2014 9:43:54 AM UTC-5, Lim wrote:
> Dear all,
>
> I would like to ask if someone know a code to calculate a 2D Pearson correlation as:
>
>
>
> r^2=(Sum wi*(Mi-M)*(Oi-O))^2 /((Sum wi*(Mi-M)^2)*(Sum wi*(Oi-O)^2))
>
>
>
> Sum runs from i=1 to N. N is the total number of grid cells.
>
> Mi and Oi are the values in the grid cell i and wi is a normalized weight (area) of grid cell i. Sum wi=1 (Sum from i=1 to N).

>
>
>
> IDL has C_Correlate and R_correlate but none of them include the wi factor.
>
>
>
> I will appreciate any assistance.
>
>
>
> Lim

Subject: Re: 2D Pearson correlation coefficient
Posted by David Fanning on Fri, 31 Jan 2014 13:12:37 GMT
View Forum Message <> Reply to Message

rryan@stsci.edu writes:

> You should read a few blogs (including David's) on "how to ask a help question." I truly mean no disrespect.

Here is the link. I've had the feeling lately we are not posting this often enough:

http://www.catb.org/~esr/faqs/smart-questions.html

Cheers,

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

Subject: Re: 2D Pearson correlation coefficient
Posted by Phillip Bitzer on Fri, 31 Jan 2014 16:11:57 GMT
View Forum Message <> Reply to Message

OK, a few things.

1) That is not a Pearson correlation. Coefficient of determination, maybe.

2) Like Craig said, if you don't know anything about what the weights should be, you should use

one. You *could* weight by some knowledge of the errors in the data. We don't know what they are, so we can't tell you what to use. Besides, you're the person closest to the data. You would know best.

3) But, you better have good reason for weighting the data. I strongly suggest you curl up with Bevington's Data Reduction book before you go manipulating the data.

4) Matt told you how to do this. But here goes anyway:

```
;get some data
m = RANDOMU(seed, 100)
o = RANDOMU(seeed2, 100)

wi1 = 1 ;assuming no points are "more important" than others
wi = (randomn(1l, 100)+2) > 0 ;Just making up *something* for the weights
wi /= TOTAL(wi) ;make sure weights add to one

mMean = MEAN(m)
oMean = MEAN(o)

r2 =  TOTAL(wi1*(m-mMean)*(o-oMean))/TOTAL(wi1*(m-mMean)^2)/TOTAL( wi1*(o-oMean)^2)
```

Of course, you should make this a function, with appropriate parameters, and then try it with arbitrary weights.
