Subject: Re: Statistics : T-test, P-value
Posted by JMB on Tue, 15 Jan 2008 11:21:23 GMT
View Forum Message <> Reply to Message

>> I came to know these correlation by using 'CORRELATE'
>> function. However, I want to know whether these
>> correlation is reasonable or not. So I want to calculate
>> p-value by T-test. Is there any idea calculate p-value
>> through t-test through IDL?

Hi Nick,

I don't know if you can find something useful for you in the following
small program:
After calculation of the Pearson correlation coefficient with the IDL
correlate function,
you can test your coefficient in 2 ways:

- by computing its CONFIDENCE INTERVAL based on the number of data
points.
0 should not be included in this confidence interval to claim that
your correlation is significant.

- by using a t-test of a null hypothesis on the correlation
coefficient

```
;------------------------------------------------------------ ----------------
; Significance Tests on Pearson's Correlation
; Based on http://davidmlane.com/hyperstat/B62223.html

PRO CORR_TTEST,corr=corr,N=N,sl=sl, ro=ro

; corr: Pearson's correlation coefficient to be tested
; N:   Number of samples
; sl: Significance level accepted (ex:0.05,0.001,)
; ro: Correlation value predicted by theory (Null Hypothesis)

; Assumptions
; 1. The N pairs of scores are sampled randomly and independently.
; 2. The distribution of the two variables is bivariate normal.

; NULL hypothesis is ro=ro

IF N_Elements(ro) EQ 0 THEN ro=0
IF corr EQ 1.0  THEN corr=0.9999999d  ; avoid Floating divide by 0
IF ro  EQ 1.0  THEN ro=0.9999999d  ; avoid Floating divide by 0
IF corr EQ -1.0 THEN corr=-0.9999999d  ; avoid Floating divide by 0
IF ro  EQ -1.0 THEN ro=-0.9999999d  ; avoid Floating divide by 0
```

```
 ;------------------------------------------------------- ----------------
; COMPUTE CONFIDENCE INTERVAL OF CORRELATION COEFFICIENT
 ;------------------------------------------------------- ----------------
; Conversion of Pearson's correlation to the normally distributed
variable zp
; Fisher's transformation

zp=0.5*alog((1+corr)/(1-corr))

;Compute zp standard error

sig_zp=1/sqrt(N-3)

; Compute z value from significance level sl
; 99% confidence interval example corresponds sl=0.01 and gives to
z=2.58

z=gauss_cvf((sl)/2.)

low_zp=zp-z*sig_zp
high_zp=zp+z*sig_zp

r_high=(exp(2*high_zp)-1)/(exp(2*high_zp)+1)
r_low=(exp(2*low_zp)-1)/(exp(2*low_zp)+1)

print,''
print, "High End Case for r: ",r_high,"     Low End Case for r:
",r_low

; Preliminary result of significance based on Pearson correlation
interval
; If the 0 is included in the range between r_low and r_high,
; You can't claim your result is Statistically significant at
significance level (sl)
; (or confidence level (1-sl)

IF r_low LT 0 AND r_high GT 0 THEN $
print,"This is NOT a statistically significant relationship!" ELSE $
print,"This is a statistically significant relationship!"
 ;------------------------------------------------------- ----------------

 ;------------------------------------------------------- ----------------
; T test significance
 ;------------------------------------------------------- ----------------
print,''
print,'---T TEST result
```

```
; if Null hypothesis is ro=0

IF ro EQ 0 THEN BEGIN

Df=N-2
t=corr*sqrt(Df)/sqrt(1-corr^2)
pt=2*(1-T_PDF(t, Df))

IF pt LT sl THEN $
print,"The correlation is significant repect to significance level
",string(sl,format='(F7.5)') ELSE $
print,"The correlation is NOT significant repect to significance level
",string(sl,format='(F7.5)')

ENDIF ELSE BEGIN

; if Null hypothesis is ro<>0

zpro=0.5*alog((1+ro)/(1-ro))
zt=(zp-zpro)/sig_zp
pzt=2*(1-GAUSS_PDF(zt))

IF pzt LT sl THEN $
print,"The null hypothesis that the population correlation is
",string(ro,format='(F7.4)')," can be rejected." ELSE $
print,"The null hypothesis that the population correlation is
",string(ro,format='(F7.4)')," CAN'T be rejected."

ENDELSE

END
 ;------------------------------------------------------------ ----------
```

Regards,
Let us know,

Jérôme