相关

协方差：$SP = \sum XY - \frac{\sum X \sum Y}{n}$

Pearson 相关：$r = \frac{SP}{\sqrt{SS_X SS_Y}}$

参数

① 相关的显著性检验. $\boxed{t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}}$ $t_{crit} = t_{0.05}(df = n-2)$

比较两个 $r_1 r_2$ 是否有显著差异. 查表得 Zr_1 Zr_2 或

$Z_r = 5[\ln(1+r) - \ln(1-r)]r$

$Z = \frac{Zr_1 - Zr_2}{\sqrt{1/(n-3)}}$

X	X^2	Y	Y^2	X·Y
等	×	等	○	—
	×		○	—
级	×	级	○	—
$\sum X$	$\sum X^2$	$\sum Y$	$\sum Y^2$	$\sum XY$

Spearman 相关

非参. 1、排序. 分等级

2、计算 SS_X SS_Y 与 SP $\quad r_s = SP / \sqrt{SS_X SS_Y}$ Person

$D = Y_{等级} - X_{等级}$. $r_s = 1 - 6\sum D^2 / [n(n^2-1)]$ spearman 数据不能重复.

4. 进行查表进行显著性检验. 或者 ①

点二列相关.

是非. $r_{pb} = \frac{\bar{X}_p - \bar{X}_q}{S_X}\sqrt{pq}$ 与 t 之间关系：$r_{pb}^2 = \frac{t^2}{t^2 + df}$ $df = n_1 + n_2 - 2$

进行显著性检验用公式 ①

Kendall 和谐系数

等级 $W = \frac{\sum R_i^2 - \frac{(\sum R_i)^2}{N}}{\frac{1}{12}K^2(N^3 - N)}$

$r_s = \frac{KW-1}{K-1}$ (spearman)

		N 个事物				
		1	2	3	4 …	N
K 个评定者	1	×	×	×	× …	×
	2	×	×	×	× …	×
	3	×	×	×	× …	×
	⋮	⋮	⋮	⋮	⋮	⋮
	k	×	×	×	× …	×
	R_i					$\sum R_i =$
	R_i^2					$\sum R_i^2 =$

回归初步

斜率 $= b = \frac{SP}{SS_X}$ $\quad$ 截距 $= a = \bar{Y} - b\bar{X}$

$SP = \sum XY - \frac{\sum X \sum Y}{n}$ $\quad SS_X = \sum X^2 - \frac{(\sum X)^2}{n}$

$\hat{Y} = \bar{Y} - b\bar{X}$

估计的标准误的计算步骤. 估计的标准误 $= \sqrt{\frac{SS_{error}}{df}} = \sqrt{\frac{\sum(Y-\hat{Y})^2}{n-2}}$

又∵ $SS_{error} = (1-r^2)SS_Y$

∴ 估计的标准误 $= \sqrt{\frac{(1-r^2)SS_Y}{n-2}}$

χ^2 匹配度检验.

$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$ $\quad f_o$ = 观察值 $\quad f_e$ = 期望值

$df = C - 1$ $\quad \chi^2_{crit} = \chi^2_{0.05}(df)$ $\quad \chi^2_{crit} \leq \chi^2_{ob}$ 拒绝 H_0

若题目中给出的是正态分布. 求出每组上限下限对应的 z 值, 得出 z 值所对应的概率. $z_1 = \frac{X_1 - \bar{X}}{S}$ $z_2 = \frac{X_2 - \bar{X}}{S}$ $df = K - 3$

χ^2 独立性检验.

		因素 B					
		1	2	3	…	C	行和
因素 A	1	×	×	×		×	m_1
	2	×	×	×		×	m_2
	…						
	r						m_r
	列和	n_1	n_2	n_3		n_c	N

期望值 $= f_{rs} = \frac{m_r n_c}{N}$

$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$

$df = (R-1)(C-1)$

$\chi^2_{crit} = \chi^2_{0.05}(df)$

χ^2 检验的效应水平

$\Phi = \sqrt{\frac{\chi^2}{N \times df_{min}}}$

$df_{min} = \min\{r-1, s-1\}$

df_{min} ↑ $\Phi_高$ ↓

χ^2 检验正态分布的 CI

$$\left(\frac{(n-1)s^2}{\chi^2_{2,0.95}}, \frac{(n-1)S^2}{\chi^2_{2,0.05}}\right)$$

非参数检验. 若有相同的方法处理秩.

1. 曼-惠特尼 U 检验

① 小 排列. 找点数. 或进行计算 $U_A = n_A n_B + \frac{n_A(n_A+1)}{2}$ $\quad U_B = n_A \times n_B - U_A$

$U = \min\{U_A, U_B\}$ $\quad U_{crit} = U_{0.05}(n_A, n_B)$

$U_{ob} < U_{crit}$ Rj H_0

② 大 两个样本中至少有一个容量大于 20

$\mu = \frac{n_A n_B}{2}$ $\quad \sigma = \sqrt{\frac{n_A n_B(n_A + n_B + 1)}{12}}$ $\quad z = \frac{U - \mu}{\sigma}$

$|z_{obs}| > z_{crit}$ rj H_0.

2. 符号检验法. (相关样本) 前后对比的正负号

小 $(n<25)$ 在符号检验表中, 直接将较少的符号的数目与临界值进行比较. $obs < crit$ Rj H_0.

大 $(n>25)$ $\mu = np \; (p = \frac{1}{2})$ $\quad \sigma = \sqrt{npq}$ $\quad z_{obs} = \frac{r - \mu}{\sigma}$

z_{obs} 和 z_{crit} Rj H_0

3. 维尔克松 T 检验. (相关样本) 前后数值之差绝对值的顺序

小 $(n<25)$ $R_- = X_1$ $\quad R_+ = X_2$ $\quad T = \min\{R_+, R_-\}$ $\quad T_{crit} = T_{0.05}(n)$

$T_{obs} < T_{crit}$ Rj H_0

大 $(n>25)$ $\mu = \frac{n(n+1)}{4}$ $\quad \sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$ $\quad z = \frac{T - \mu}{\sigma}$

4. 克-瓦氏单向方差分析 (多顺序独立样本)

小 排列之后. 计算 $\sum R_A$, $\sum R_B$, $\sum R_C$.

$H = \frac{12}{N(N+1)}\sum\frac{R^2}{n} - 3(N+1)$ 查表 $\alpha = 0.01$

$H_{obs} > H_{crit}$ Rej H_0.

大 $(k>3$ 或 $n>25)$ 查自由度为 $k-1$ 的 χ^2 分布.

5. 弗里德曼双向方差分析. (多顺序相关样本)

小

评价者	原始分数 (A	B	C) k	秩次 A	B	C
1				1	2	3
2				1	3	2
3						
(n)						
				R_A	R_B	R_C

$\chi_r^2 = \frac{12}{nk(k+1)}\sum R^2 - 3n(k+1)$

$\chi^2_{crit} = \chi^2_{r(0.05)}(k, n)$

若 $\chi^2_{obs} > \chi^2_{crit}$

则 Rej H_0. 弗里德曼表

大 查自由度为 $k-1$ 的 χ^2 分布表.

Pvalue = power = Pr $\Phi\left(Z_\alpha + \frac{|\mu_0 - \mu_1|}{\sigma/\sqrt{n}}\right)$ one side

$= \Phi\left[Z_{\alpha/2} + \frac{|\mu_0 - \mu_1|}{\sigma/\sqrt{n}}\right]$ two side

1. 对两个独立样本的假设检验

① 独立样本 t 检验. F值检验.

$S_1^2=\frac{SS_1}{df_1}$　$S_2^2=\frac{SS_2}{df_2}$　$S_P^2=\frac{SS_1+SS_2}{df_1+df_2}$　S_P

$S_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{S_P^2}{n_1}+\frac{S_P^2}{n_2}}$　$t_{obs}=\frac{\bar{x}_1-\bar{x}_2}{S_{\bar{x}_1-\bar{x}_2}}>t_{0.05/2}$ Rej H_0

检验效应 $ES=\frac{\bar{x}_1-\bar{x}_2}{S_P}$　Cohen's d

② 相关样本 t 检验.

$SS_D=\sum D^2-\frac{(\sum D)^2}{n}$　$S_D^2=\frac{SS_D}{n-1}$　$S_{\bar{D}}=\sqrt{\frac{S_D^2}{n}}$

$t_{obs}=\frac{\bar{D}-\mu_D}{S_{\bar{D}}}$　$t_{obs}>t_{0.05/2}$　Rej H_0.

检验效应　$ES_D=\frac{\bar{D}}{S_D}$

基本公式汇总.

$SS=\sum(X-\mu)^2=\sum X^2-\frac{(\sum X)^2}{N}$　$\sigma=\sqrt{\frac{SS}{N}}$　$\sigma^2=\frac{SS}{N}$

对样本：$S^2=\frac{SS}{n-1}$　$S=\sqrt{\frac{SS}{n-1}}$

人数的计算

$$n=\frac{\sigma^2(Z_{1-\beta}+Z_{1-\alpha/2})^2}{(\mu_0-\mu_1)^2}$$

单因素和重复测量方差分析

1. 单因素.

方法A	方法B	方法C
0	0	0
0	0	0
0	0	0
0	0	0
T_1	T_2	T_3
SS_1	SS_2	SS_3
n_1	n_2	n_3
$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$

G

$SS_{总和}=\sum X^2-(G^2/N)$　$SS_{总和}=SS_{组间}+SS_{组内}$

$SS_{组内}=SS_1+SS_2+SS_3$　$SS_{组间}=\sum T^2/n_i-G^2/N$

$MS_{组间}=SS_{组间}/df_{组间}$　$MS_{组内}=SS_{组内}/df_{组内}$

$F_{obs}=MS_{组间}/MS_{组内}$

事后检验　HSD检验

$HSD=q\sqrt{MS_{组内}/n}$

$\bar{x}_1-\bar{x}_2>HSD$

$F(df_{组间}/df_{组内})0.05=F_{crit}$.

Scheffe 检验.　$MS_{组内}$ 整体　$df_{组间}$ 整体

$SS_{组间}$ 比较两组间　$MS_{组间}=\frac{SS_{组间}}{df_{组间}}$

$MS_{组内}$

$F_{obs}=MS_{组间}/MS_{组内}$

2. 重复测量方差分析.

被试	练习次数 1	2	3	K
A	0	0	0	P_1
B	0	0	0	P_2
C	0	0	0	P_3
D	0	0	0	P_4
n	T_1	T_2	T_3	N
	SS_1	SS_2	SS_3	

$df_{组间}=K-1$　$df_{组内}=N-k$

$df_{被试间}=n-1$　$df_{误差}=df_{组内}-df_{被试间}$

$SS_{总和}=SS_{组间}+SS_{组内}$

$SS_{组内}=SS_1+SS_2+SS_3$

$SS_{组内}=SS_{被试间}+SS_{误差}$

$SS_{被试间}=\sum(P^2/k)-G^2/N$

$SS_{误差}=SS_{组内}-SS_{被试间}$

$MS_{组间}=SS_{组间}/df_{组间}$　$MS_{误差}=SS_{误差}/df_{误差}$

$F_{obs}=MS_{组间}/MS_{误差}$

$F_{crit}=F_{0.05}(df_{被试间},df_{误差})$

进行 HSD 事后检验

方差分析的效应大小.　$f=\sqrt{F/N}$　0.01　0.25　0.40

和总体均值相关的估计

1. 方差已知情况下总体均值的估计.

① 点估计　$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$

② 区间估计　$\left[\bar{x}-\frac{Z_0\sigma_0}{\sqrt{n}},\ \bar{x}+\frac{Z_0\sigma_0}{\sqrt{n}}\right]$

2. 方差未知情况下总体均值的估计.

② 区间估计　$\left[\bar{x}-t_0\sqrt{\frac{SS}{(n-1)n}},\ \bar{x}+t_0\sqrt{\frac{SS}{n(n-1)}}\right]$

3. 独立组总体均值差异的估计

① 点估计　$\mu_A-\mu_B$

② 区间估计　$\left[\bar{X}-\bar{Y}\pm t_0\sqrt{\frac{SS_A+SS_B}{n_A+n_B-2}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}\right]$

4. 相关组总体均值差异的估计

① 点估计　$\bar{D}=\bar{X}-\bar{Y}$

② 区间估计 $\left[\bar{D}-t_0\sqrt{\frac{SS_D}{(n-1)n}},\ \bar{D}+t_0\sqrt{\frac{SS_D}{(n-1)n}}\right]$

二因素方差分析

	B_1	B_2
A_1	0 0 0 0	0 0 0 0
A_2	0 0 0 0	0 0 0 0

因素A (a=2)	因素B (b=2)		
	A_1B_1	A_1B_2	A_1
	A_2B_1	A_2B_2	A_2
	B_1	B_2	G

$a=b=2$

A_1B_1

X	X^2
0	0
0	0
0	0
0	0
0	0
$\sum X$	$\sum X^2$

$SS_{A_1B_1}=\sum X^2-(\sum X)^2/n$

$A_1B_1=\sum X$

$A_1=A_1B_1+A_1B_2$　$A_2=A_2B_1+A_2B_2$

$B_1=A_1B_1+A_2B_1$　$B_2=A_1B_2+A_2B_2$

$G=A_1+A_2=B_1+B_2$

$df_A=a-1$　$df_B=b-1$

$df_{A\times B}=(a-1)(b-1)$

$df_{处理内}=N-a\times b$

$SS_{总和}=\sum X^2-\frac{G^2}{N}$

$SS_{处理间}=\sum\frac{AxBy^2}{B}-\frac{G^2}{N}=\frac{A_1B_1^2}{n}+\frac{A_1B_2^2}{n}+\frac{A_2B_1^2}{n}+\frac{A_2B_2^2}{n}-\frac{G^2}{N}$

$SS_{处理内}=\sum SS_{AxBy}$

$SS_A=\sum\frac{A_x^2}{bn}-\frac{G^2}{N}$　$SS_B=\sum\frac{By^2}{an}-\frac{G^2}{N}$

$SS_{A\times B}=SS_{处理间}-SS_A-SS_B$

$MS_{处理间内}=\frac{SS_{处理内}}{df_{处理内}}$　$MS_A=\frac{SS_A}{df_A}$　$MS_B=\frac{SS_B}{df_B}$

$MS_{A\times B}=\frac{SS_{A\times B}}{df_{A\times B}}$　$F_A=\frac{MS_A}{MS_{处理内}}$　$F=\frac{MS_B}{MS_{处理内}}$

$F_{A\times B}=\frac{MS_{A\times B}}{MS_{处理内}}$　$F_{crit}=F_{0.05}(df_{A\times B},df_{处理内})$

$HSD=q\sqrt{MS_{组内}/n}$