

广义估计方程

按随机抽样的原则，回归分析要求每个观测对象是独立的，互相之间没有相关性。但在随访研究中，同一观测对象有多次观测，这些记录有内部相关性。在家系研究中，来自同一家庭的成员间有相关性。这样的数据也不满足广义线性模型条件。Liang and Zeger (1986)提出广义估计方程（GEE）分析有内部相关的数据。

研究对象编号（ID）

研究对象编号（ID）是用于确定来自同一研究对象的变量。

相关结构的类型

来自同一研究对象的数据内部有相关性，但是什么样的相关性呢？常见的相关结构的类型是独立、互换、自回归、非结构化。

应变变量分布和关联函数

因变量 Y 的分布和联系函数和广义线性模型相同。

该模块调用 GEE，对一个或多个应变变量进行回归分析，自变量类型不限，可以引进交互作用项。

例 1: demo 数据中 FMYID 表示核心家系编号，SUBJ 表示个人编号，同一 FMYID 内可有多多个 SUBJ，他们之间有内部相关性，调用 GEE 分析 SBP。输入界面如下：

广义估计方程

标题: 广义估计方程

选择分析对象: 所有数据记录

应变变量

变量名	分布类型	联系函数
Systolic BP, mmhg	Gaussian	Identity

自变量

变量	选择
Age, years	.
SEX	.
Height, m	.
Weight, kg	.
Education	.
Occupation	.
SMOKE	.

分层变量

自动检验与选择的自变量(S)的交互作用

如用GEE:

研究对象编号: fmyid

内部相关类型: independence

刷新 保存 查看结果

输出结果

广义估计方程

结局变量: Systolic BP, mmhg

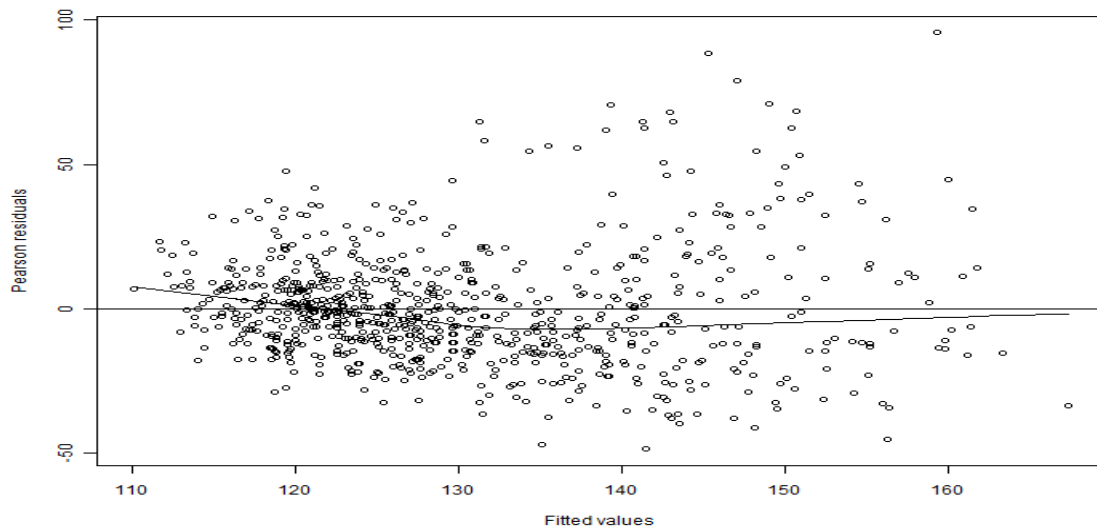
变量分布与联系函数: gaussian

模型: $SBP \sim AGE + SEX + HEIGHT + WEIGHT + \text{factor}(EDU.NEW) + OCCU.NEW + SMOKE$

	Estimate	Std.err	Wald	Pr(> W)	95%CI low	95%CI upp
(Intercept)	152.891	27.4012	31.1334	0	99.1847	206.5972
AGE	0.8073	0.0691	136.5497	0	0.6719	0.9427
SEX	-6.5256	2.5057	6.7825	0.0092	-11.4367	-1.6144
HEIGHT	-32.6217	18.6239	3.0681	0.0798	-69.1245	3.8811
WEIGHT	0.3026	0.1494	4.1	0.0429	0.0097	0.5955
factor(EDU.NEW) 2	0.1125	2.0149	0.0031	0.9555	-3.8367	4.0616
factor(EDU.NEW) 3	0.5628	2.068	0.0741	0.7855	-3.4904	4.616
OCCU.NEW	-4.4351	1.4496	9.3605	0.0022	-7.2763	-1.5938
SMOKE	-3.6017	2.1987	2.6836	0.1014	-7.9111	0.7076

Number of cluster:	195
Maximum cluster size:	9
Correlation structure:	independence
Number of observations used:	787

模型诊断: 自动绘制残差与预测值散点图, 以判断残差的方差是否独立于预测值。



例 2: 如在例 1, 引进两变量交互作用项于模型中, 输入界面如下:

广义估计方程 ?

标题: 广义估计方程

选择分析对象: 所有数据记录

应变变量

变量名	分布类型	联系函数
Systolic BP, mmhg	Gaussian	Identity

自变量

变量	选择
Age, years	S
Education	.
Occupation	.
SMOKE	.
Body mass index, kg/m2	S

分层变量

sex

自动检验与选择的自变量(S)的交互作用

如用GEE:

研究对象编号: fmyid

内部相关类型: independence

刷新 保存 查看结果

注：在自变量框内右击变量名选“修改选择”，以启动“自动检验分层变量与被选择的自变量的交互作用”功能。分析结果首先是按分层变量的分层分析，然后是合计的分析及交互作用检验，如下：

结局变量: Systolic BP, mmhg
 变量分布与联系函数: gaussian
 模型: SBP~AGE+factor(EDU.NEW)+OCCU.NEW+SMOKE+BMI

亚组分析: SEX = Male

	Estimate	Std.err	Wald	Pr(> W)	95%CI low	95%CI upp
(Intercept)	84.5583	13.932	36.8373	0	57.2516	111.8649
AGE	0.8215	0.1225	44.9892	0	0.5815	1.0616
factor(EDU.NEW) 2	3.7041	3.173	1.3627	0.2431	-2.5151	9.9233
factor(EDU.NEW) 3	1.9478	3.0773	0.4007	0.5268	-4.0836	7.9792
OCCU.NEW	-6.6122	2.3311	8.0456	0.0046	-11.1812	-2.0432
SMOKE	0.5266	2.3826	0.0489	0.8251	-4.1433	5.1966

BMI	1.0526	0.5573	3.5669	0.0589	-0.0398	2.1449
-----	--------	--------	--------	--------	---------	--------

Number of cluster:	175
Maximum cluster size:	5
Correlation structure:	independence
Number of observations used:	395

亚组分析: SEX = Female

	Estimate	Std.err	Wald	Pr(> W)	95%CI low	95%CI upp
(Intercept)	93.407	10.7856	75.0017	0	72.2673	114.5468
AGE	0.8439	0.0834	102.413	0	0.6804	1.0073
factor(EDU.NEW) 2	-0.1972	2.4595	0.0064	0.9361	-5.0177	4.6233
factor(EDU.NEW) 3	4.3393	2.7085	2.5667	0.1091	-0.9694	9.648
OCCU.NEW	-2.3653	1.9254	1.509	0.2193	-6.1391	1.4086
SMOKE	-2.0004	2.2587	0.7844	0.3758	-6.4274	2.4266
BMI	0.3917	0.4728	0.6863	0.4074	-0.535	1.3184

Number of cluster:	172
Maximum cluster size:	5
Correlation structure:	independence
Number of observations used:	392

合计:

	Estimate	Std.err	Wald	Pr(> W)	95%CI low	95%CI upp
(Intercept)	93.9165	9.1522	105.3021	0	75.9783	111.8548
AGE	0.8175	0.0694	138.9667	0	0.6816	0.9535
factor(EDU.NEW) 2	0.0977	2.014	0.0024	0.9613	-3.8498	4.0451
factor(EDU.NEW) 3	0.4057	2.0669	0.0385	0.8444	-3.6455	4.4569
OCCU.NEW	-4.5945	1.4481	10.0663	0.0015	-7.4329	-1.7562
SMOKE	-3.8009	2.1997	2.9858	0.084	-8.1123	0.5105
BMI	0.7674	0.3907	3.859	0.0495	0.0017	1.5331
factor(SEX) 2	-5.3674	2.1152	6.4389	0.0112	-9.5133	-1.2215

Number of cluster:	195
Maximum cluster size:	9
Correlation structure:	independence
Number of observations used:	787

分组因素与选择的自变量交互作用分析: SEX with Age, years Body mass index, kg/m2

	Estimate	Std.err	Wald	Pr(> W)	95%CI low	95%CI upp
(Intercept)	114.5706	10.1126	128.3574	0	94.7499	134.3913
AGE	0.6945	0.0921	56.8379	0	0.514	0.8751
factor(EDU.NEW) 2	1.2352	2.1829	0.3202	0.5715	-3.0434	5.5137
factor(EDU.NEW) 3	0.7709	2.0651	0.1394	0.7089	-3.2766	4.8184
OCCU.NEW	-3.8097	1.4004	7.4005	0.0065	-6.5546	-1.0649
SMOKE	-3.3621	2.2159	2.3021	0.1292	-7.7052	0.9811
BMI	-0.0918	0.4542	0.0408	0.8399	-0.982	0.7985
factor(SEX) 2	-45.9877	15.0177	9.3773	0.0022	-75.4224	-16.553
AGE:factor(SEX) 2	0.2779	0.1448	3.6857	0.0549	-0.0058	0.5617
BMI:factor(SEX) 2	1.4489	0.6549	4.8942	0.0269	0.1652	2.7325

Number of cluster:	195
Maximum cluster size:	9
Correlation structure:	independence
Number of observations used:	787

上表中 AGE:factor(SEX) 2、BMI:factor(SEX) 2 分别是 AGE、BMI 与 SEX 的交互作用项。