

## 实验四：一元线性回归

实验目的：掌握一元线性回归的基本步骤，了解逐步回归等变量筛选方法及含义，了解线性诊断及模型评估的基本方法。能正确解读分析结果，规范书写检验报告。

**关键：**选用合适的检验方法，正确解读分析结果

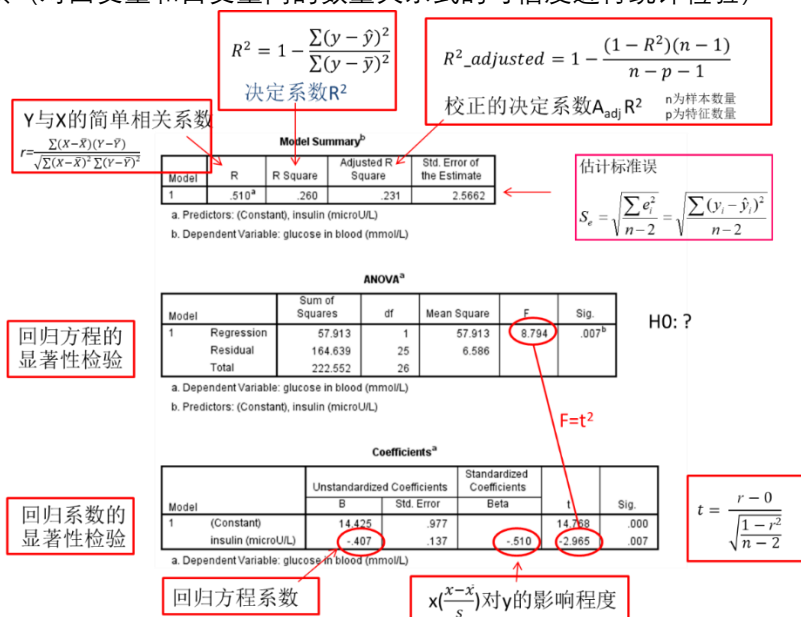
### 1. 线性回归

#### 1.1 一元线性回归（简单回归：一个因变量，一个自变量）

27 名糖尿病人的血清总胆固醇、甘油三脂、空腹胰岛素、糖化血红蛋白、空腹血糖的测量值列于下表中，试建立血糖与胰岛素的一元线性回归方程。

序号 $i$	性别	总胆固醇 (mmol/L)	甘油三脂 (mmol/L)	胰岛素 ( $\mu$ U/ml)	糖化血 红蛋白(%)	血糖 (mmol/L)
		$X_1$	$X_2$	$X_3$	$X_4$	$Y$
1	男	5.68	1.90	4.53	8.2	11.2
2	男	3.79	1.64	7.32	6.9	8.8
3	男	6.02	3.56	6.95	10.8	12.3
4	男	4.85	1.07	5.88	8.3	11.6
5	男	4.60	2.32	4.05	7.5	13.4
6	男	6.05	0.64	1.42	13.6	18.3
7	男	4.90	8.50	12.60	8.5	11.1
8	男	7.08	3.00	6.75	11.5	12.1
9	男	3.85	2.11	16.28	7.9	9.6
10	男	4.65	0.63	6.59	7.1	8.4
11	男	4.59	1.97	3.61	8.7	9.3
12	男	4.29	1.97	6.61	7.8	10.6
13	男	7.97	1.93	7.57	9.9	8.4
14	女	6.19	1.18	1.42	6.9	9.6
15	女	6.13	2.06	10.35	10.5	10.9
16	女	5.71	1.78	8.53	8.0	10.1
17	女	6.40	2.40	4.53	10.3	14.8
18	女	6.06	3.67	12.79	7.1	9.1
19	女	5.09	1.03	2.53	8.9	10.8
20	女	6.13	1.71	5.28	9.9	10.2
21	女	5.78	3.36	2.96	8.0	13.6
22	女	5.43	1.13	4.31	11.3	14.9
23	女	6.50	6.21	3.47	12.3	16.0
24	女	7.98	7.92	3.37	9.8	13.2
25	女	11.54	10.89	1.20	10.5	20.0
26	女	5.84	0.92	8.61	6.4	13.3
27	女	3.84	1.20	6.45	9.6	10.4

- 1) 将数据复制到 SPSS，修改一下变量的名称等属性。
- 2) 做血糖与胰岛素的散点图，查看是否具有相关性
- 3) 建立回归方程 SPSS 步骤：分析 → 回归 → 线性 → 因变量：血糖，自变量：胰岛素 → 确定
- 4) 结果解读（对因变量和自变量间的数量关系式的可信度进行统计检验）



5) 书写回归报告

根据变量间的散点图，探究因变量与自变量之间是否可能具有相关性

利用 SPSS 建立因变量与自变量的一元回归方程，并对回归方程的可信度进行统计检验

$H_0: \beta = 0$

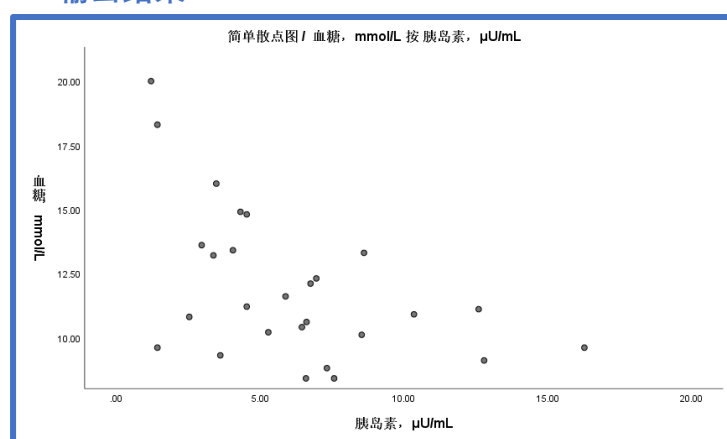
$H_1: \beta \neq 0$

$\alpha = 0.05$

检验方法：一元线性回归检验

操作步骤：略

输出结果：



先绘制散点图，发现血糖与胰岛素具有负相关关系，且可能具有线性关系，可以进行回归分析。

模型摘要 <sup>b</sup>				
模型	R	R 方	调整后 R 方	标准估算的错误
1	.510 <sup>a</sup>	.260	.231	2.56623
a. 预测变量: (常量), 胰岛素, μU/mL				
b. 因变量: 血糖, mmol/L				

ANOVA <sup>a</sup>						
模型		平方和	自由度	均方	F	显著性
1	回归	57.913	1	57.913	8.794	.007 <sup>b</sup>
	残差	164.639	25	6.586		
	总计	222.552	26			
a. 因变量: 血糖, mmol/L						
b. 预测变量: (常量), 胰岛素, μU/mL						

系数 <sup>a</sup>						
模型		未标准化系数		标准化系数	t	显著性
		B	标准错误	Beta		
1	(常量)	14.425	.977		14.768	.000
	胰岛素, μU/mL	-.407	.137	-.510	-2.965	.007
a. 因变量: 血糖, mmol/L						

检验结论：由于 p=0.007<0.05，应当拒绝 H0，接受 H1，即生物学意义上，认为血糖关于胰岛素的线性相关关系显著。

根据系数表格，得到具体的回归方程为：

$$\hat{y} = -0.510 x + 14.425$$