excel回归分析（回归之后的理解）

2020年11月26日

16:26

<https://zhuanlan.zhihu.com/p/57875710>

知乎老哥对于excel进行回归分析的方法步骤

计算机生成了可选文字:
103
99．5
105
98
103．5
103
104
103
277
257
255
278
306
268
285
286
272
285
MARYOIJTPIJT
回归统计
MultipleR
RSquare
StedRSq
标准误差
观测值
方差分析
回归分析
残差
总计
〔L875648
J76676
〔L737605
7．713788
10
1
8
9
Coefficient.
Intercept一2〔L2585
XV引b-2．907618
SS
1564．88
476．0203
204J9
《标准误差
57．99625
〔L566976
1564．88
59．50253
tStat
一〔L34931
5128292
F
26．29938
P-value
〔L735875
〔L000898
nificanceF
〔L000898
Lower95Uer95下限95℃上限95℃
一153．998113．4811一153．998113．4811《
1．6001694．2150661．6001694．215066
=2．907618×一2〔L2585

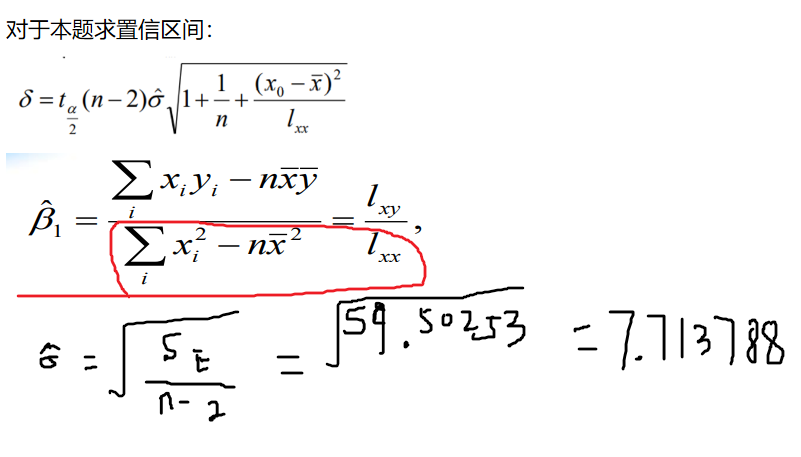
例如上题：

表1：

1. 回归统计表，Multiple R即相关系数R的值，大于0.8表示强正相关。
2. 回归统计表中的R Square是R平方值，R平方即R的平方，又可以叫判定系数、拟合优度，取值范围是[0,1],R平方值越大，表示模型拟合的越好。一般大于70%就算拟合的不错，60%以下的就需要修正模型了。这个案例里R平方0.76676，相当不错。
3. Adjusted R是调整后的R方，这个值是用来修正因自变量个数增加而导致模型拟合效果过高的情况，多用于衡量多重线性回归

表2：

1. 第二张表，方差分析表，df是自由度，SS是平方和，MS是均方，F是F统计量，Significance F是回归方程总体的显著性检验，其中我们主要关注F检验的结果，即Significance F值，F检验主要是检验因变量与自变量之间的线性关系是否显著，用线性模型来描述他们之间的关系是否恰当，越小越显著。这个案例里F值很小，说明因变量与自变量之间显著。（线性回归即假设H0:X之前的系数a=0，p值小于阿尔法值，拒绝原假设，即表明是线性相关的，相关方程显著）
2. 残差是实际值与预测值之间的差，残差图用于回归诊断，回归模型在理想条件下的残差图是服从正态分布的。
3. 第三张表我们重点关注P-value，也就是P值，用来检验回归方程系数的显著性，又叫T检验，T检验看P值，是在显著性水平α（常用取值0.01或0.05）下F的临界值，一般以此来衡量检验结果是否具有显著性，如果P值>0.05，则结果不具有显著的统计学意义，如果0.01<P值<0.05，则结果具有显著的统计学意义，如果P<=0.01，则结果具有极其显著的统计学意义。T检验是看某一个自变量对于因变量的线性显著性，如果该自变量不显著，则可以从模型中剔除。



**如何求置信区间：**

计算机生成了可选文字:
来
回
源
归
平方和
自由度
丿一1
。一2
均方和
壕：斗/
F比

计算机生成了可选文字:
SR：0一2：〗0一到]2：属2巛

计算机生成了可选文字:
斗：0'！一）2，其自由度为：”一2
从而有ST：斗+

计算机生成了可选文字:
ST：一2

计算机生成了可选文字:
/一2）都是0无偏估计，

置信区间：

计算机生成了可选文字:
Yo的概率为1一的预测区间为：[

计算机生成了可选文字:
：0仂一2卮]+一+一№二型《

其中：（注意lxx的求法）

计算机生成了可选文字:
一理父Y

这里的β1即x的系数。

多元线性回归分析：

<http://blog.sina.com.cn/s/blog_a20c88b601014j9x.html>

计算机生成了可选文字:
与成矩阵形式如下
与成矩阵形式为．
则系数求解公式如下：w=(XTX)-lXTY
其中，x丆代表×的转置矩阵，()l、X)-中的一1代表矩阵的逆矩阵。