# Quantitative Methods

Covariance:协方差，两个变量**的总体**[**误差**](https://baike.baidu.com/item/%E8%AF%AF%E5%B7%AE/738024)，这与只表示一个变量误差的[方差](https://baike.baidu.com/item/%E6%96%B9%E5%B7%AE/3108412)不同。

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps1.jpg

n: sample size

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps2.jpg: ith observation on variable X

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps3.jpg: mean of the variable X observations

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps4.jpg: ith observation on variable Y

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps5.jpg: mean of the variable Y observations

Correlation coefficient:相关系数，在1和-1之间。1和-1代表完全在一条直线上。

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps6.jpg

C:\Users\Li\AppData\Local\Temp\ksohtml16636\wps7.jpg

检验相关系数的显著性，使用t检验（样本数小于30且方差未知）。t的计算使用相关系数r。C:\Users\lz\AppData\Local\Temp\enhtmlclip\Image(1).png，t落在置信区间之外则拒绝。

检验相关系数的显著性，使用z检验（样本数大于30）做平均值的差异性检验。

Z=\frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}

检验线性回归中所有斜率相关系数均为0，用F检验。

分子是自变量对因变量变化决定比例，SSE代表除自变量之外对因变量的变化决定因素。F越大代表回归模型越可靠。

回归线是估计值和真实值之间差的平方的和最小的线。

**斜率系数b1等于，截距用平均值加上斜率来求解。**

变化程度的决定系数：

衡量因变量的总体变化程度，真实Y值和平均值之间的差的平方和。和方差不尽相同，**方差variance=SST/(n-1)；total sum of squares**

预测值和平均值之间的差的平方和, 衡量由独立变量所揭示的**因变量的变化程度**。

真实值和预测值之间的差的平方和。**因变量中未被解释的部分**。

Standard error of estimates:

检验斜率系数是否等于某个假设值：

给定的点估计，n-2自由度下的检验统计量是：

Reject if

拒绝假设代表着斜率系数与假设值不一样。

有时候不会直接报是否拒绝关于回归参数的特定假设，他们会报告p-value。这个p-value是空假设能被拒绝的最小显著性水平。

这个和上面那个t检验的公式其实是一样的。。。

或者计算系数的关键值区域：

If the estimated value falls into this interval, you can’t reject the hypothesis of estimated value.

Image(13)

The lower the p-value reported for a test, the more significant the result.

Adjusted 不会因为加了自变量就变大。

is nondecreasing in the number of independent variables, so it is less reliable as a measure of goodness of fit in a regression with more than 1 independent variable than in a one independent variable regression.

调整过的r平方肯定比原始的小，因为原始的会随着k增加而放大的，但调整的不会。