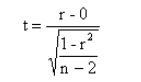


作业3:

R^2很高时，假设检验一定会是显著的吗？

从P值看假设检验显著时，R^2一定高吗？

T-stats和R^2的关系：



R够高，t越大，p-value越小。

数学问题，用数学公式，不要拍脑袋呜呜QUQ。

但是反过来就不一定了，因为这个公式里还有n这个变量。

R=0，t=0;

在一元回归中F检验和T检验等价。

R^2衡量了模型的拟合优度，而假设检验(t检验)则是从参数是否显著的角度出发（看x对y 是否有明显的影响)。

所以R^2的假设检验的p值没有必然关系。

通过写公式推导：

作业4:

为什么数据被复制了一份，对假设检验的p值有影响，对假设检验的哪里有影响呢？

H\_hat公式（XTX）^{-1}Xy里与数据量是无关的。n增大时，t的自由度会增加；所以n会影响我的整个p-value,同样的数据n越大，p会变小，不significant会significant，所以不能数据不够复制数据，这虽然对estimation不影响，但是会导致test不准。

样本量越大，显著性水平需要降低设置(变严格)。

样本量变大,p值变小，线性回归拟合的值不会变。

这题不是特别理解，猜测是跟中心极限定理大样本量类似的思想有关。样本量越大，我提出的原假设在更大样本量的情况下，由p值的变小说明我提出的原假设更有力度，即偶然性，拒绝原假设的概率会变小。

在数据复制一份的情况下，样本方差也会变小。

