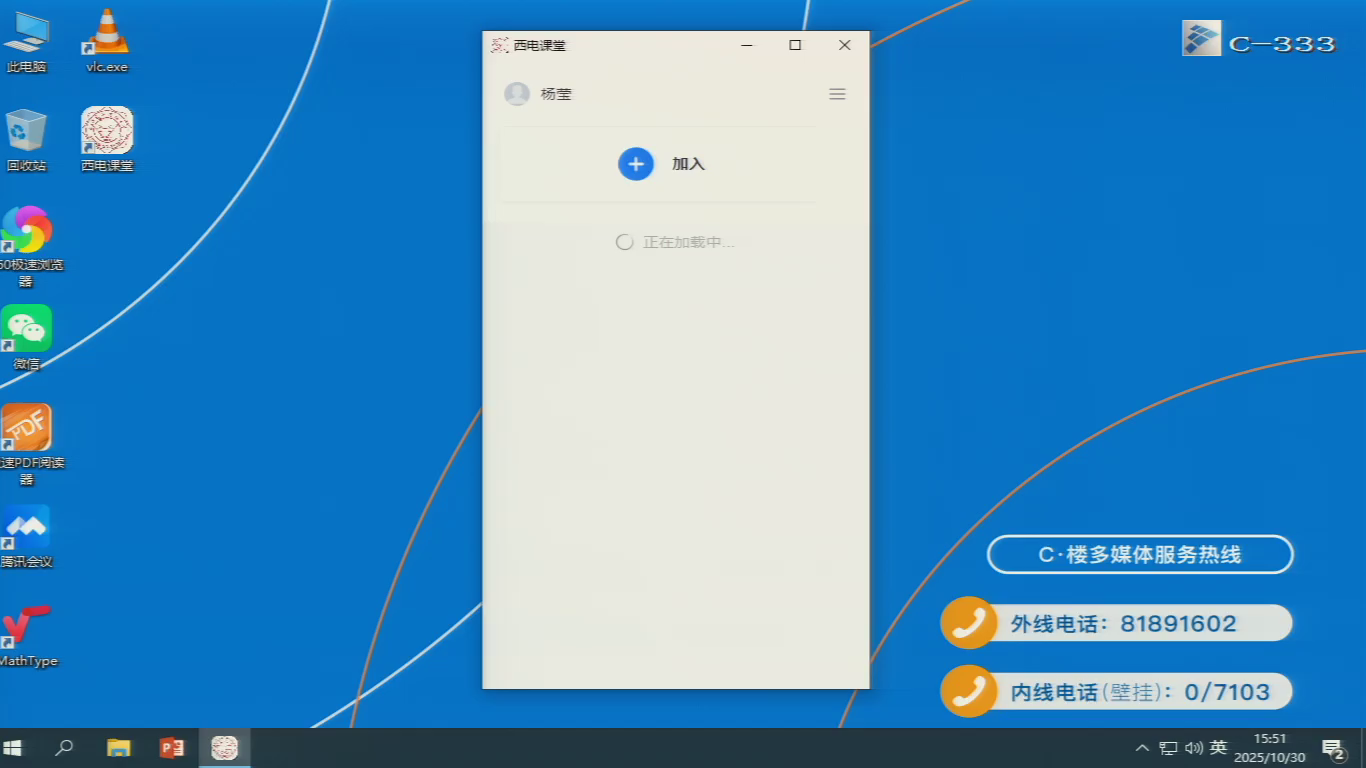
EM114026大数据计量经济分析2025年10月30日第9周星期四第7-8节-pptVideo\_原文

2026年01月04日 12:57



发言人 00:01

就是。















发言人 06:24

好的，那我们这节课继续我们的这个第三章经典线性回归模型的这样一个课程的学习。在上一节课我们是不是给大家讲的这个我们献祭回归当我们把这个什么你估计出来的，或者说你用这个OLS估计出来的这样一个结果，我们还要保证它的什么。除了我们通过OLS估计出来的一个什么最优的参数值，也就是说它的什么它的贝塔的取值，我们已经得到了最优的这样一个估计。

发言人 07:08

贝塔容貌，贝塔一帽。这个贝塔零号和贝塔一帽，我们再确定它的这个估计出来假设图，估计出来它的值，也估计出来它的这个正负号。那这个就是我们带到什么OIS，我们估计的这个公式里面，我们一算就能算出来对吧？

发言人 07:31

但我算出来的这个值，假设说贝塔一等于什么0.6，因为0.6或者0.4什么的，0.3什么的，就是我们还要确定的一个是什么，我得到的这个数我们还要确定他的什么，我找个同学来回忆一下，上节课一直在强调的一个问题是什么？我们知道了这个大小，也知道了这个方向，那我们下一步要确定的是什么呢？张子贺。能回忆的起来，我们上节课讲的啥？就发生在昨前天，对吧？我们估计出来的这个0.6，它是一个正数，然后大小是0.6，X对Y的影响是X变动一个单位，Y变动0.6个单位，对吧？

发言人 08:27

那下一步我想要的结论，那我下一步再确定的是什么呢？检验显著性检验对吧？就是说虽然我现在知道了X对外的影响是正向的，而且大小是0.6。但是我还要确定的一个是什么？我得到的这个估计的什么参数，它是一个随机变量，这个随机变量他不应该？有可能会取到0对吧？

发言人 09:03

好，好，请坐。这个就是我们上节课一直给大家讲的，我们想要确定X对Y的影响，我们最终想得到什么？X对Y具有显著性的影响，这个显著性的影响啥意思？就是不可能出现X对Y没有影响，我也可以access这样一个结果的情况。所以你说在贝塔一这样一个关键参数的它的什么变动的可能性里面，我们不应该包含理由，或者说我们应该要确定贝塔一这个值它显著区别于0，这个时候我们才能得到说什么X对Y具有显著性的影响。

发言人 09:51

这个就是我们上节课给大家讲的那在不在做这样一个显著性检验的时候，我们大概有哪几个步骤呢？郭子轩。就是我要检验这个参数，它是不是什么显著区别于0，我要做哪几个步骤呢？做假设？假设检验也就是说设这个H0H？

发言人 10:32

就是我要构造一个统计量，它是什么？比如说类似服从已知的一些分布，比如说T分布，正态分布什么对吧？那我怎么构造这个统计量呢？是不是就根据我们上上节课讲的这个，我们在这个线性回归模型里我们有一系列的假设，对吧？我们有假设这个蛮缪什么它是服从正态分布的，所以这个贝塔它也是服从正态分布的那这样我们可以根据这个什么贝塔减去增值，我假设它的增值是0，它的真值是零，也就意味着我假设对于总体来讲，X对Y是没有影响的对吧？那这样我们去构造一个什么贝塔减去增值，我假设它增值为零，然后再除以它的标准差，这个时候它就是一个T分数了，对吧？那如果样本量很大，它是不是就是一个正态分布了？

发言人 11:33

然后我们再看看吧，我们构造的这个统计量，它在这个T分布或者正态分布里的哪个位置，我把这个数据带进去，估计出来的贝塔以及估计出来的什么。是不是还有估计出来的贝塔的这个什么标准差，对吧？是不是我们就可以得到这样一个显著性的检验？

发言人 12:09

我们的估计量减去我们的假设值，再除以估计量的这个标准差，那我们去检验的时候，好我把系统计量也构造出来，然后再下一步是什么呢？对，就是我要去看这个T值是不是超过了我的这个什么临界值，对吧？比如正态分布里面，它小概率小概率分布在两边，它就是什么可能落在我们的拒绝域里面，对吧？那如果落在拒绝域域里面的时候，我们应该怎么拒绝袁雅芝，也就是得到一个什么结论？就不等于0，对吧？那不等于零也就意味着。就是贝塔一不等于0，也就意味着什么？

发言人 13:10

贝塔一是不是代表的是X对Y的影响，如果是贝塔一不等于0，那也就意味着X对Y是有肯定有影响的对吧？好，请坐。这个就是我们在进行这个假设检验的时候，我们进行的几个主要的步骤。

发言人 13:30

除了刚刚这位同学讲的，我们可以用这个什么T检验，我们还可以用什么检验呢？就除了我们比较它的这个T值和临界值，我还有什么方法呢？于博陈。就除了我比较我算出来这个分布里面的T值，我看它是不是落在这个是不是比这个尺寸大，这更小对吧？这有几个曲子，1.682.58，那我比较完了，我还有没有其他的方式来看一下，我这个是不是显著的呢？

发言人 14:17

还有什么方法？这个图是什么图？是不是概率密度图对吧？那是不是相当于这个一减A或者说一减去这个自信水平。他就是我们认为通常都会发都可能发生的事情，对吧？那我们是不是也可以算，如果说我们把这个值T值代进去，我们是不是也可以求出这个面积来，对吧？那我们把这个面积求出来是不是，也可以看一下这个面积大小，然后来决定我们的这个什么，我们这个是不是落在这个什么接受欲或者落在拒绝域里面。

发言人 15:07

除了T检验或者说是这个T值的比较，我们是不是也给大家讲的清楚。P值的检验也就是说它干它它的那个T值对应过来不就是它的那个概率，就他的P值P就是把这个他把这个单词的字母提出来，它就把它叫做是P值。好，请坐。那我们在进行这个P值检验的时候，我们怎么样来判断它是一个他是一个显著性的影响的？李恩泽。

发言人 15:47

就是T检验我就比较临界值对吧？如果它大于比如说2.58，那我就认为它落在拒绝域里边，就是P值越大越好。那P值它是越大越好。

发言人 16:16

P值是越大越好吗？应该是越小越好，对吧？越小越好就意味着什么？如果P越接近于0，也就越是一个不可能的事件。因为P值是吗？P值代表了我们中间它这个事件发生的概率，因为它是真实发生的，但是你算出来它竟然是一个小概率事件。那这个时候也就说明肯定不是现实错了，肯定是你的原假设错了，对吧？那这个时候你就可以去拒绝掉我们的原假设。

发言人 16:53

好，请坐。大家一定要注意到，我们去判断X对Y有出现影响的时候，如果你用的P值，那这个P值一般比如说它越接近于0，你越能判断它是一个什么具有显著性的影响的这个情况。那这个时候你就有理由去拒绝调研。

发言人 17:15

假设说你算出来的P值，我假设说你算出来P值是0.9的话，那0.9也就为了也就意味着中间大家看这个分布图里面和概率布图里面这一段白色的面积它是0.9。那如果说我把实际数据带进去，我算在实际的这个白色的区域是0.9的面积。那那也就是说明什么？这是一个很正常的事情，对吧？也是说也就说它是一个可以接受的结果，那这个时候也就说明你没办法去拒绝掉原假设了，对吧？你没办法去拒绝掉原假设，你没办法得到这个贝塔一不等于零的这样一个结论了。

发言人 17:58

这个是我们在进行P检验的时候，它的步骤也是一样的，只是区别于那什么，我们最后是依赖于它，相当于是一体两面对吧？它如果确定了相应的T值，你就可以根据那个T值两边围起来的面积，你就可以算出来那个probability，就是它的概率P值，那也同样的这个得到的结果也是一样的。所以说很多时候我们在这个手算的时候，你可能能算出来气质。因为你带进去这个贝塔的这个估计值，再除以它的标准差，你就能算出来T值。

发言人 18:43

手算可以，但是通常我们在进行这个分析的时候，现在一般大家都很容易用到软件。那用软件的话，你通常就不用再去什么把那个T值它算出来之后，你再去对比你这个书本后面的这个什么T表再来看了。他一般也会把这个什么它那个T值对应出来的这个P值，就它的这个概率值直接给到你这边。你直接看到这个P值的数据，比如说它是0.05，你就可以知道它是不是有显著性的影响，对吧？

发言人 19:20

那假设说我们现在得到的一个P值是什么？0.05，那这个时候我们认为它是显著的还是不显著的呢？杨佳好，因为它是显著的，是显著的对吧？因为它小于那个什么，那你给如果说我给定的这个什么置信水平是0.01的话，那这个时候它还是显著的。他是不是还依赖于你的船？置信水平对吧？你说你的阿法就当你的这个P值小于你的R法才能拒绝掉原假设。虽然通常来讲我们这个up取0.05就可以，对吧？

发言人 20:05

好，请坐。那这个就是我们在进行假设检验，或者说去判断一个变量对另一个变量显著性影响的时候，它几种不同的方法。虽然听起来是不同的方法，但实际上只是那个概率密度图里面，你看的是那个那个轨迹上点的值，还是包围起来的那个面积值，但它俩代表的含义是一样。同样的我们来看一下这个例子，就是我们把这个数值代进去的时候，你就可以计算出来。

发言人 20:47

除了我们这个P值检验，我们在上一节是不是还讲了有其他的方法？朱成瑞。除了我们这个P值的跟临界值比较，除了这个P值和显著性水平进行比较，我还有什么方法能看出来这个X和Y有显著性的影响关系呢？

发言人 21:32

能想起来吗？

发言人 21:39

这什么想不起来了。好，请坐。何志广。你被你说他感染了是吧？我再看心意。

发言人 22:06

好，你已经预习了，后面。好，请坐。王静芬。对，我们还可以看，因为它是一个什么随机的变量，我们能得到的这个0.6是它的均值？那我们加减一个分布，也就Z2分之2乘以它的标准差，是不是我们可以得到一个假设说我这个案子取的是什么？假如我案子取的是0.05的话，我是不是可以算出来一个贝塔一均值是0.6，然后95%可能出现的这个什么区间我们可以算出来，对吧？就是说95%的可能性。

发言人 23:00

这个贝塔一是在这个均值为0.6，然后左右两边是什么？0.6减去Z2分之200乘以它的S1，也是它乘以标准差。这边就是0.6加上这个Z2分之28乘以它的S1，你说他95%的可能性出现在这个区间。那这个区间里面，我怎么样来判断它是显著的呢？是我看这个区间的大小吗？还是他的这个什么？那我怎么看呢？我已经把这个区间算出来了，那就我就直接给定了。

发言人 23:44

假设如果这个0.6加减的话，假设我这个Z2分之2乘以S一算出来是假如说是0.7，那我这边就是-0.1，这边就是0.6加0.7就是1.3就说他有这个贝塔一，它有95%的可能性取-0.1到1.3，那均值只是说均值是0.6。那那正常分布里面就说它很大的概率在0.6上面，但是95%的可能性在这里。因为它取这些值都是可以接受的，或者说可能出现的取值，对吧？

发言人 24:23

那现在这个是什么？从-0.1到1.3都是可能的，那这个时候你能得到这个X对Y的影响是显著的吗？可以的对吧？对，好，行，那你就被我带偏了，看一下有没有。就一定要提防对吧？王泽西这个时候我得到是-0.1到1.3，这个时候我能得到X对Y的影响是显著的吗？它均值也挺大的，0.6对吧？那也是个正数。为什么呢？

发言人 25:10

因为我刚把另外一个同学带听了，他他说可以，对吧？杨启超可以吗？虽然现在已经大家都知道是不能得到显著性的结论了，但是为什么呢？就是-0.1到1.3都是可取的。现在资金区间算出来是这个，我们怎么样通过这个区间来看，我们不能判断X对Y有显著性的影响。

发言人 25:58

怎么看？对，好，请坐。伊斯拉伊利。叶潇然。说要通过一下体检，我们还要再通过体检。你就不想看这个资金区间，我是知道的对吧？好，请坐。上节课我们讲了？老师一直强调我们要判断这个贝塔一是不是显著，或者说X对Y是不是有显著性的影响。其实最主要的是看什么，他是不是跟零存在显著性的区别，因为它是个随机变量。

发言人 26:47

如果说大家可以看到，这个-0.1到1.3是95%的这个区间里面有这9，这个是什么？负零点这边是一点，也就意味什么？它包含了零对吧？也是零也在他的什么可能出现的取值范围内。而且它不是这个零取值，它不是一个小概率事件，对吧？

发言人 27:14

它在这个95%的可能性的这个区间里面，那这个时候，我们就不能得到的结论是非常一显著区别于0。因为它也有可能取到0的值。虽然说我们看起来0.6是一个正向的，而且它还是一个挺大的数值。但是有的时候你可以看到，如果说我的这个S1，也就是它的标准差比较大的时候，即使我得到的这个估计参数是一个比较大大的正数。但是因为我这个标准差算出来它的什么经过了这个零的取值和贝塔一取0也是可能出现的估计结果。而且这个估计结果也是可取的那这个时候我们就不能得到这个X对Y影响就是显著的这样一个结论。

发言人 28:09

但是如果说我们算出来这个资金区间，假设说这个Z2分之2，乘以这个S1，最后算出来是什么0.4。那这边就是0.2，这边就是1.0对吧？那这个时候大家可以看到，我90%多的可能性，基本上就是便利了贝塔一所有可能性的取值，对吧？它是从0.2到1.0的都是正数，也都是啊什么大于零的取值。

发言人 28:44

那这个时候大家可以看到这个置信区间里面它就不出现0。不出现0，那这个时候我们就可以看到贝塔一取不到0，就它是一个水晶变量，它也取不到0。那这个时候我们就可以说明这个X对Y的它是一个显著性的影响。这个就是我们上节课给大家讲的三种不同的，我们来判断X对Y是不是存在显著性影响。这种检验的方法，它的逻辑其实都是类似的，对吧？

发言人 29:19

我们在构造这个置信区间的时候，我们也要看它的这个什么分布的你构造的这个T分布，或者说是什么正态分布，它的这个临界值的取值。比如说你要娶一个T2分之2或者说Z2分之2来进行替代。那这个时候你就可以算出这中间这个微塔一可能的取值。那这样我们就可以看到它判断这个零的取值在不在我们这个置信区间。

发言人 29:56

置信区间叫什么？那就是你可信的区间，一个可取的区间里面，这个可取的区间里面，或者说你相信的这个区间里面，它它不存在0。那这个时候你就可以去拒绝掉这个原假设，然后得到什么，得到我们这个X对Y存在显著性影响的这样一个结果。这个后面除了我们在检验一个模型，或者说我们估计出来的这个OOS这个结果之外，我们除了对X和Y的这个系数进行检验，我们也要对这个码模型整体的表现进行检验，对吧？所以说我们通过这个什么T检验，假使说，我们得到了这个斜率贝塔1以及截距上。

发言人 30:48

贝塔理论在统计上都是显著的。贝塔理论也是类似的。比如说我们很多时候不关注这个，很少会讨论这个杯子，也是为什么？因为我们我们或者说我们进行分析的时候，很多人都是关注变量跟变量之间的互相影响关系，对吧？但因为贝塔零它是一个截距项，你如果是啊零，那就说明你构造的这个模型可能在截距项的有问题，那其他的他也不说明什么。

发言人 31:23

东西很多时候，所以我们讨论的时候都是着重讨论这款变量前面的这个系数被采集。我们通过什么T检验，我们可以发现这个嘛斜距和斜率都是统计上显著的那这个时候只能说明什么？我们估计出来的这个样本回归线跟我们一直看的那个图，那个蓝线他能很好的去拟合我们那些点那些样本数据，对吧？但是我们是不是也有一个问题呢？我们想要得到的结论是关于总体的结论，或者说是我们这个什么真实值距离，我们这个什么回归线，要么回归线它是不是也有一定的距离？那究竟这个距离好坏，就说这个距离的什么大小，是不是它也有一定的意义，对吧？那这个时候，我们就可以用什么构造这样一个准则来度量我们这个样本回归线他对我们这个实际数据拟合的好坏。也就是说我们现在虽然找到了这样一条最好的线。

发言人 32:45

那这样一条最好的线，比如说我在这组数据里面表现出来和我用另外一个数据，你也拟合出来一条最好的线，那这两条线可能也有区别对吧？比如说他对这个数据的代表性也可能会存在区别。这个区别我们用来表针就用这样一个礼盒优度，这样一个这样这样这样这样一个表示。热力和优度它也称为判定系数。

发言人 33:16

通常我们用这个R光来表示，也是说regression的这个平方，它是用来度量Y的数据变差中能被X解释的比例。那也就意味着虽然我们现在得到了X对Y这样一个影响，它是显著的对X对外存在影响。但是这个影响或者说X我们现在只能得到X的变动，会引起Y的显著性的变动，但是这个变动它引起的什么能解释的范围，或者说Y的变动里面有多少是因为X的变动而产生的，我们是不是也需要去看一下对吧？那这个时候我们用图里面来看，就可以很容易的看到。

发言人 34:09

我们从这个法离差平方和这样一个分解里面，我们就可以看到总的离差真实值减去真实值，不真实值减去总体的这个估计值是Y，对吧？Y8是我们的均值，它分为两个部分，一个部分是什么？真实值，距离我们这个样本回归线的差值。这个上节课之前我们也给大家讲过，对吧？这个是我们的残差值。还有一个部分是什么？就是我们的样本回归线的预测值减去我们的这个Y的均值，它也有一个YI mom的这样一个曲子。所以我们能估计出来两个离差平方和就是Y2加上我们残差的估计值，这个表示就是总离差值。

发言人 35:11

通过我们把这个什么离差平方和进行分解，大家可以看到离差平方和我的真实值减去我的什么总体的什么均值，那这样我们把它分解成加减一个什么Y帽，你把它带进去，就可以分解成一个残差平方和的形式，加上这个Y2减去Y82Y8平方和的这样一个形式。那通过这样一个离差平方和，或者说是用这样一个什么平平方差的这样一个形式来解释什么被解释变量YI它的什么样本观测值和均值的差平方和均值大家可以想到我60码代表我的这个平均性的规律，对吧？就总体的这个离差平方和或者是总变差和就表示为什么我们的每一个观测值Y减去Y均值和YR平方求和，它是一个总变差或者说总的离差平方和，比如我们用TSS来表示。

发言人 36:36

这个它是什么？这个。总的变差。那TSS表示的是吗？它表示的是被解释变量的样本估计值和均值的这样一个离差平方和对吧？它就是Y一个减去Y把那个平方这个是什么？就是由回归解释的这个平方和用这个EXS来表示。也就是说什么？大家可以看到我这个变动总变动，比如说总变动就真实值减去你的均值，对吧？

发言人 37:17

总变动里面有一部分可以被解释的，可以被解释出来，就是我们已经勘察到了这个规律了，就是这条线，那这条线距离我的这个均值这部分，是可以被我们回归所解释的。但是真实值距离什么样，本核对线上面的那部分，它就称为残差平方和。残差平方和就是不被这个回归模型解释的。所以大家可以看到我们通常这个墙它就放在这个后面一个庙里面一个随机朗动的部分，它不被我们这个模型解释的部分，就是在上面这个红色的部分。

发言人 37:56

那大家从这个图里面是不是就可以看到我一个真实值距离平均值的什么所有的信息差，或者说是所有的信息总和，它就可以分成两个部分。一部分是绿色的部分，这个绿色的部分就是我们模型已经估计出来这个参数了，对吧？所以这个部分其实就是确定的，你可以算出来的，所以它也这样的可解释的离差平方和。那不能被解释的，就是在我们这个线之上的，就真实值减去我们这个样本回归线的残差的平方和，这个部分我们就用RSS来表示。所以如果我们这样一看就可以看到我们总的离差平方和TS的是不是就等于这两部分求和，对吧？比如说ESS加上这个RSS，那我们的拟合优度是不是这样你就可以去构造出来了？我把这个平方和可以解释的部分占什么，占这个总的部分，我是不是就可以算出来这个模型我能解释的这个变动范围有多大，或者变动能够有多大？所以我们构造的这样一个尺码拟合优度R方，它的度量就很简单。

发言人 39:23

就是说我们看这个绿色的部分占整个黄色部分的比例，我是不是就可以知道我这个模型到底性能好坏了，那这个时候我们构造这个R方就可以写成它ESS除以TSS对吧。也就是说我们这个图里面绿色的这个平方和除以什么黄色的这个平方和，那是不是也就可以写成什么，是不是就可以写成一减去这个红色的部分平方和除以黄色的和平方和，对吧？因为它们求和起来是一，我把它相减的话是不是就可以算了？因为为什么要这样写？因为我这个残差平方和真实的数据万一我们是有的，然后估计出来这个样本回归线，你把每一个X值代进去，它的预测值YY帽子不是也可以算出来吗？是不是每一个I的那个长差一I你是不是也就算出来了？这个在之前的课程里我们也算过，对吧？

发言人 40:30

那这个时候我用一减去一平方求和，再除以我的总离差平方和，总离差平方和我也很好算，因为我真实值是知道的，然后外把比如说我均值我也知道，所以这样一算，我其实就可以算出来我这个你和优度R方的那大家从这个式子里我们就可以看到这个拟合优度它的含义就表示什么，它表示的是什么因变量观测值。它总变异部分或者它这个Y的变动当中，它由模型所解释的部分，总共占多少的比例，对吧？也就说它这个Y的变动能够被我们这个模型来解释。我们这个模型来解释什么？如果你只放了一个变量，那就是X来解释。如果你放了两个变量，那就是Y的变动。总共可以由X1X2来解释它的这个程度到底是多少，对吧？所以大家可以看到我们这样一个礼盒优度的界定。

发言人 41:40

下面就可以发现这个R方有一个性质？它的取值肯定是什么？是不是它是一个0到1的值对吧？因为它是一个比值，解释程度的多少就是这个绿色的部分占黄色部分的大小。我们来看这个模型的性能好坏，或者说这个模型能解释Y的变动多少。在在这样的一个定义下，这个R方的取值它就是从0到1的那大家可以看到我这个R是R方是啊越大越好，还是越小越好呢？

发言人 42:30

何志广。

发言人 42:37

为什么呢？

发言人 42:48

他是不是这个绿色的部分占黄色部分的比例，对吧？那如果说他他越大越好，就意味着么他越接近是吧？接近一对吧？也就是意味着这个真实，也就是说他这个样本回归线离这个真实值越近对吧？或者说它这个模型可以解释的部分就越大对吧？就是绿色的部分越大好几个。所以我们通常去判断这个R方，我们就会认为如果这个R方它越接近于一，它就说明我们这个样本回归线，它对这个观测值的拟合程度越好。

发言人 43:36

大家可以想象到一个极端的情况是吧？一个极端的情况是我刚好就收集了这样一种它的点，就是这样的这样一组数据。当我去回归这个样本回归线的时候，我就刚好就能回归或者说拟合出来这样一条线的话，那我画的不直，但是实际上这条直线它是不是就完全经过了我们收集到的这些样本数据的？那这个时候，大家可以看到我的这个码，这个图里面的残差值是不是就是零了，因为我这一条线完全解释了现实的变动，你说这些点全部可以完全的由我这条线来解释，就不存在我解释不了的这个情况。那这个时候说明你构造这个模型，或者说你这个模型里面的这个变量对这个Y的影响已经全部被你捕捉到了，对吧？那这个时候就说明什么，我们这个样本移位线呢，它对我们观测值的拟合程度就越好。如果说这个R方的取值越接近于0，那就说明这个样本回归线对这个观测值的拟合程度就越差。也就是说他就没办法被你这个模型，就是Y的变动没办法被你这个模型来解释。

发言人 45:19

因为这个R方它里面存在一个什么我们存在一个什么自变量个数的问题。大家就会发现，如果说我们在这个方程里面增加一个自变量，它通常不会改变总体差平方和。也就是说我们这个黄色的部分，它不随着这个X就是我这个X变量的数量的变动。因为什么？

发言人 45:51

它是真实值距离我这个Y均值的差距，对吧？你说大家可以看到它它跟X的影响是没有关系的。但是有一个问题是什么呢？当我们这个X我的这个变量X越多的情况下，它通常会什么增加这个RSX也就是说即使你加进来的是一个没有什么用的X它也会增加你这个模型的可解释性。也就是说它有一个，如果你得到的这个R方比较小，你不断的往这个模型里面增加变量，它都会增加R方的取。那这样就会造成一个，我想得到如果说我单纯用这个R方来判别一个模型的好坏的话，那我可能得到的一个结论，我往里面加越多的X它就越好，吧？那这显然不是一个什么不是一个科学的事情。

发言人 46:57

所以我们为了避免他从在这样一种R方天然是跟随着X的变量增加而增加的这样一个性质，我们就对进行一个修正，也就说让我这个牙膏不水，或者说不太受我这个X自变量数量的影响。那这个时候我们就构造这样一个判定系数，这样一个修正的这样一个R方，来去代表我们这个拟合优度来使什么用这个什么这样一个模型。自变量次数过多，我们同样这个自由度会降低的这样一个问题，我们就对它进行一个修正。修正的方法就是什么？我们把它的RSS和TSS的自由度做一个平均。

发言人 47:55

也就是说我表示的什么，每一个自由度里面包含的就是什么离差的大小，那TSS我什么是什么YI减去Y均值对吧？这个自由度是多少？我们是不是之前给大家讲过。我们找一个同学来回忆一下。

发言人 48:26

伊斯拉伊利，所以我说他这个点名不是随机的，它这个YI减去Y8均值，它的自由度是多少？就是总共我们的这个I有N个变量，这个时候我的均值是固定的那这个时候我的这个自由度是多少？

发言人 48:58

不知道。好，请坐。邓宇。它自由度是多少？自由度其实就可以解释成我们这个变动的范围。如果形象一点来讲的话，我给N个数据，假设说这里是N个数据，这N个数据里面我均值已经知道了，或者说我这个均值已经知道是0.6的那这个时候我这里面的这些数据我可以任意取N个数吗？

发言人 49:40

可以吗？大家可以想象到，如果第一个数任意取，第二个数任意取，第三个数任意取，取到第N减一的数据。你说N减1，前面那个数据我全部固定好了，那是不是第二个数据我还能自由取吗？不行，为什么呢？因为我还要保证它的均值是0.6，对吧？好，请坐。

发言人 50:07

所以这里就存在一个问题是吗？当我们这个均值确定了之后，他就会什么损失一个自由度。也就是说当你把前面N减一个数据自由取了之后，管你取什么值。比如说你取23500，管你取什么值，你最后一个数字应该是给定的对吧？因为你最后一个数字要保证你的这个均值是等于0.6，也就说最后一个数你就不能再任意取了。所以它能自由取的数其实就是什么N减一个，这个就是自由度的概念。很多人到了硕士博士都分不清这个自由度是什么意思，他可能老是这样，一跟你讲你就明白了。

发言人 50:54

那同样的RSS里面，它的自由度是吗？就是N减K减1，这个减一同样也是我们均值固定了，它就少掉一个。但是K是什么？在最开始我们讲这个模型Y等于贝塔0加贝塔一X这个X我们是不是也给大家讲过它什么固定的去，所以它固定了几个变量？我们能取的数就越少，或者说这个人自由变动的范围也就越少。所以它会是什么N减去K个自变量，再减去一个一也做均值固定在我的这个影响。那TSS是总的，所以它这个就是YI距离均值，它的这个TN减一的这样一个自由度就可以了。

发言人 51:47

我们再把RSS和TSS的这两个这两个表示带进去了，大家就可以发现我是不是就可以写成这个什么，相当于就R方的基础上再乘以一个什么自由度的这个比例的调整。这样我这个K的变化的影响，因为我自变量的影响就不会影响这个礼盒优度的这样一个变化。那这样这样一个修正的修正之后的这个判定系数，我们就用这个什么R8方，就是我们在R方上面加一层，这个就表示我们修正之后的判定系数。

发言人 52:34

它通常用在什么？用在一元线性回归模型还是多元线性回归模型。郭子俊。多元对吧？因为它只有当X自变量比较多的情况才会出现这个问题对吧？如果是只有一个自变量，那大家可以看到它带进去这个修正之后和修正前是一样的。

发言人 53:03

好，请坐。这个就是我们在进行什么多元回归的时候，大家一定要注意的一个问题。当我这个X变量比较多的时候，如果你想要比较两个不同的模型，它的可解释性或者说拟合优度的好坏，我们通常就用这个修正之后的这个判定系数来进行比较。

发言人 53:29

那大家可以看到我调整之后的这个R方用在多元回归里面也就意味什么？我们这个K自变量的个数是大于一的对吧？那当我这个K大于一的时候，它就有一个性质，我调整之后这个R不R方肯定是小于R方的对吧？就从这个公式里面就可以看出来，它是一个牙缝更小的数值。而且它随着这个自变量个数的增加，我这个调整的这样一个判定系数，它就越是小于我们这个R方。

发言人 54:11

第二个是我们这个调整的R方，它也是小于等于一的。但是它有一个情况，因为它的最大值它可能取到1，但是它的最小值大家可以看到，我剪完之后，它的这个最小值它就不再是零了。也就是它的这个可能取值的范围是一个比零更大一点的值到1。这个只是说当判定系数R方比较小的时候，我这个调整之后的R方甚至有可能是负值。也就是说可能是什么一个比零更小的值。

发言人 54:54

那那同样的我们用之前的这样一个样本数据，把它带到我们这个R方的这样一个判别式里面。我们就可以是不是就可以算出来我们这个判定系数R方的取值，比如说我们用之前的这个例子，残差平方和我们通过什么构造的这个样本回归线OAS估计出来的贝塔零帽和贝塔一帽。然后我再把XI带进去，就可以得到YI帽的什么估计值或者说预测值，预测值再减去什么真实值，是不是就可以算出这个残差。然后再把残差每个平方项求和，是不是就可以得到这最后得到1.287。那同样的Y方，也就是说真实值减去YY8，它的这个什么做一个差值，然后再求平方也可以算出来。然后两个一代带到这个R方里面，就算出来它是什么0.9723的这样一个取值。那同样我们在算这个修正的这样一个判定系数，就可以算完是吧？0.9688。

发言人 56:09

所以我们就可以看到，当我们得到了R方这样一个取值的时候，我们就可以如果从语言上去表示，他就可以说成什么我们这个家庭消费支出是我们的Y对吧？我要看他消费的变动有多少是受这个收入的影响。就可以看到，我这个家庭消费支出的总变动里面，我们有96.88%都是由我们估计出来。一个怎样的回归模型进行了解释，这个就是我们在评判我们goole US估计出来的条样本回归线，它到底对这些点的代表代表了多少，或者它代表这些点的好坏，我们就能算出这样一个值，0.9688来表示这些点所代表这个Y的变动里面有多少是由我们这个模型来解释的那如果说我们这个模型是一个什么一元的线性回归模型，也就是我们这个模型里面只有这一个X的变量。那这个时候，你就可以说这个Y的变动有96.88%都是由X的变动所引起的。只是说很多时候我们这个模型里，它不是一个多元回归吗？你除了X一还有X2、X3、X4。比如说我们这个M41X20，它打包起来影响了什么家庭消费支出，可以对它做出的96.88%这样无变动的解释。好，那我们先休息一下。

发言人 01:04:16

我们实验课是这周末上一次，然后到课程结束上一次，还是课程结束之后周六周天一起上一次呢？啊这周末是吗？行好，那我约一下，到时候。

发言人 01:06:35

实验课就约这周末，就是早上再给我约一下他十点。好，那我们继续我们这个。在上一节课我们给大家讲，除了我们这个什么系数的检验，也就说这个贝塔检验，我们是不是还可以看整个模型的或整个估计圈回归线礼盒的好坏。除了我们可以看出来这什么X对Y是有显著性影响，包括我们可以看出来这个什么X的变动。如果说一元回归的话，我们可以看出来X的变动能解释Y变动的多少百分比。或者说如果是多元回归的话，那是X1、X2、X3、X4、X5，他们打包在一起可以解释Y变动的多少程度。这是之前我们给大家讲的不同的方面。

发言人 01:07:43

接下来我们也可以看一下，这个可能刚刚有同学已经预习到了，我们在这节课开始的时候就有同学说除了我们在做什么T检验的时候，我们还可以做什么F检验。这个F检验它是什么？它它的用途是用来干嘛？来检验我们模型当中因变量和自变量之间的线性关系，它在总体上是不是显著的，然后来进行判断。这个MT检验，同样我们就可以通过这样一个假设检验，同样来进行它检验。它逻辑都是这样的，我们构造一个假设检验HEHE，然后再构造相应的统计量。然后再看这个统计量，它在我们这个分布里面，它落在接受域还是拒绝率。但也同样的，我们可以用这个F或者P值来进行进行判断。

发言人 01:08:41

我们构造这个方程联合性显著性检验，我们的目的是什么？还在多元回归时，我们要看我们构造的这个模型里面，比如说我包括XESSCSC5这些变量里面有没有能够显著性去影响Y的变量。那这个时候或者说是我要碰到你我找的这些自变量里面，你得至少有一个是不等于零的，或者有至少有一个是有显著性影响的。这个时候你构造的整个模型是不是才具有意义，对吧？或者说才具有成立的这样一个这样这样这样一个显著性的情况。所以我们在过渡到这个F检验的时候，或者说我们去检验一个方程它到底显不显著，它的含义是跟我们检验系数的不一样的。

发言人 01:09:39

我们去检验一个方程它是不是显著的，我们就看这个方程里面有没有包含至少一个变量，它是对Y有显著性影响的。所以我们构造了H0？过道等于0，就是所有都等于0，也就说X1M2X都对Y也有影响。所以贝塔一贝塔2贝塔它都是等于零的那被则假设H一就是什么呢？H一就是你这些里面至少有一个不等于0，那这个时候我这个模型就成立了，或者我这个模型就有显著性了。我们在构造这个F统计量的时候，我们就用什么这个方程ESS，也就是说我的这个什么残差平方和除以自由度，然后再除以我们这个什么RSS，也就是说我们的这个什么。这个什么RSS，它俩一除我们这个比值，它就是服从这样一个KN减K减一这样的一个分布。

发言人 01:10:57

如果说我们得到的这个F的统计量，它大于我们在给定显著性水平阿尔法的时候，我们查这个表，如果说他比他大，比营业值大，那我们就可以拒绝原假设。不然我们就接受我们就不能拒绝原假设，因为很少会说接受人家，反正我们不能拒绝原假设，这个是我们去进行这个方程联合显著性检验的时候，大家可能会觉得有一点困惑。因为这个方程的显著性跟我这个礼盒优度有什么区别？好像他俩都是针对于就不针对于具体的系数，它是针对模型的，对吧？那通常来讲，我们这个模型它对样本的拟合优度越高，总体的我们这个模型它的这个线性关系的显著性也越强。

发言人 01:11:59

F统计量我们把它构造成什么ESS除以RSS，然后再除以它各自的这个自由度的情况下，我们可以看到是不是跟我们的R方其实比较接近吧？R方是什么？这个RSS除以TSS？那大家可以看到，他俩如果我把那个拿包的形式带进来，它就是一个什么R方除以K除以一减R方除以N减K减1。它其实就是一个什么跟什么判定系数R方它是一个同向变化的取值，尤其是什么当我们这个R方等于0的时候，我的这个F取最小值等于一。如果说我这个R方越大，我这个F的统计量也越大。

发言人 01:12:51

大家可以看到，如果说我R方就是非常完美的这个情况，我完全拟合了所有的样本点的情况下，R方是不是就取一的取。那如果取一，我这个F是不是就变成了一个分母为零的这样一个式子。那这个时候，这个F甚至就是一个无穷大的这样一个取值了。通常来讲，我们什么这个F检验，它是估计我们这个回归方程联合显著性的度量。它也是判定系数R方的这样一个显著性的检验。

发言人 01:13:32

方程显著性检验在因然回归模型里面，因为它只有一个什么自变量，它就不存在什么自变量联合显著的问题。所以它也就不需要进行什么F检验，也就是F检验跟什么T检验是不是也有关联，对吧？就是说什么呢？如果说我们是一个因缘线性回归模型，如果说我做出来这个什么贝塔一，我的这个系数是显著的那这个时候我还要做F检验吗？或者我F检验是显著的吗？黄金凤。

发言人 01:14:20

如果说我是一个因缘的线性回归模型，我T检验是显著的。但是我T检验出来这个系数X对Y是有显著性影响的那这个时候我还要做F检验吗？或者F检验是显著的吗？就不需要再做F检验了，对吧？那这个时候F检验出来它的结果肯定是什么是显著的对吧？因为我们只有一个自变量，然后也检验出来这个自变量对外有影响了。比如说我这个模型本来就只有一个字段，比如说这个模型就是显著的了，对吧？

发言人 01:14:58

好显著，大家感兴趣可以去把这个方程推导一下。就是说我们取K等于一的情况，这个F这个统计量我们把它带入进去，最后推导出来就是我们这个系统T检验。也就说一元的线性回归模型里面，T检验和FF检验它是等价的，或者说是一样的。好，我们现在其实已经评估了对吧？评估了这个系数的显著性，也评估了模型的好坏，评估了这个模型的显著性与否，那我们是不是就可以拿除了。

发言人 01:15:44

比如说我们在进行这个什么导论课程中，给大家讲了我们计量几个步骤，我们去先提出问题收集数据，然后什么构建我们的模型，然后我们去估计我们的参数估计完了之后，我们是不要把它拿来什么做一些应用，那这个时候我们就可以把这个样本的线性回归模型，拿它来进行相应的预测，也就是我们的样本回归方程。它是不是能对我们这个Y的条件期望值进行预测？它主要是取决于什么？我们这个拟合值，它是不是给定X0的时候，我们这个Y的期望。对无偏有效估计量，比如说他要满足，我们这个药不会方程估计出来的，还就是我这个什么真实值的期望，对吧？我这个估计出来的万零报它的方差，也就是我这个期望的条件期望下它的这个方差。

发言人 01:17:00

所以在这个什么一减阿尔法的这样一个置信度下，我们是不是就可以把这个方总体均值，它的这个置信区间给它算出来了，吧？那这个时候我们去给他算，那就是在T2分之2，这样它就存在两个区间？一个是什么？它就是Y02以及我的预测值加减我这个分布T2分之2，我再乘以它的标准差，对吧？就是在这一圈检验里面，我也给大家讲过这个算算的方式。

发言人 01:17:40

如果说我们满足什么我们的期望，管理就是Y0报，也就是用我们叫什么样本回归模型估计出来的这个值，Y02它跟我们真实值之间的差值是等于0。那这个时候我是不是就可以证明我们的Y零帽是Y零的一个无偏估计了，对吧？那这个公公式的推导还是大家下来感兴趣，至于去推一下，就是这个Y0减YY02减去Y0，它的期望，我们可以把什么我们这个YY要么回归函数给它带进去，对吧？然后再减去他的什么违规的这样一个表达式，最后我们就可以得到这个相应的期望值的这样一个计算。最后大家可以看到它存在一个什么等于零的这样一个偏估计量。

发言人 01:18:48

所以在什么1减200的这样一个自信度的情况下，大家可以看到我们出现了几个不同的说法对吧？一个叫自信水平，一个叫自信度，对吧？那自信水平是什么？自信水平是我们犯第一类错误的概率，就阿尔法等于30.010.1 0.000.05的这样取值的那一减阿尔法就是我们的自信度与综合刚刚讲的那个分布里面这个可信的这个区间有个置信度。那这个自信度它就是什么一减去我们距离域吧？那他两个制定水平和制定度，它恰好是一个啥一相差的值，对吧？你说一个是0.95，一个是0.05这样一个取。

发言人 01:19:39

所以我们就可以看到，个子Y零它的置信区间。我们这个Y02加减T2分之2乘以S1，它这个SY02减去Y0，它的这样一个样本材。大家看一下这个例子在书上的这个各级预测。如果说我们这个什么Y02，我估计YS估计出来的这个什么要么回归线等于0.336加上0.739乘以XI那这个时候在X0等于1.5万的时候，我这个Y礼帽的取值是多少呢？我怎么样来计算这样一个取值呢？是不是我就把这个X0的取值带入进去，我是不是就得到我这个收入水平在1.5万元的时候，我的消费是多少？对吧？我做一个预测，这个预测指的就是关于貌，那就算出来是这个1.45，也就是说我收入是1.5万的时候，我就花费支出是1.445。

发言人 01:20:57

它可以作为什么总体均值外，它的个别值在X0等于1.5万处，它的这个预测的估计值。进一步我们去计算这个挖蜊帽它的方差它是等于什么？西塔方乘以N分之1，再加上X0减XR平方除以X来访问的这样一个求解方式。这个时候我们再给它开根号，是不是就可以算出它的标准差是0.207。所以我这个标准差算到了，然后我的预测值预测的个值也算到了，是不是就可以来求解它的执行区间了，对吧？

发言人 01:21:47

我就是什么各值加减，我这个分布Z2分之2法或者说大样本的情况下，我是它是一个近似的正大评估，所以用Z2分之2法，如果说是小样本的话，那就是T分布，那就是T样本这样的。我们来求解相应的什么制定区间。那给定的显著性水平，它给定阿尔法是等于0.05的，然后再查这个T分布表，我们对照着查了相应的什么2分之200的水平，然后自由度是N减2。为什么是N减2呢？我是不是有一个自变量，然后就相当于N减K减1，你把K带进去，那就是2，那它总共有十个变量，十个变量带进去了，那就是T0.0258。

发言人 01:22:44

你查一下那个书后面的表，你就会发现是吗？2.306，那这样我们把这个什么算出来，这个个值1点45再加减T2分之200乘以标准差，是不是就得到了这个执行区间的范围？也就是说在这个什么0.945到1.45之间。所以你就说什么我们得到这个区间，它的含义就变成了什么？这个总体的均值它在95%的这个区间，那同样我们可以计算个别值的它的标准差，它只是标准差存在一个差别，对吧？我们把这个标准差算出来，同样把一个个体值按95%的执行区间也可以带入进去来进行计算，就可以得到这一个个体值，它的执行区间是在0.394到2.497之间。

发言人 01:23:54

那个值预测是我们什么？如果说我们对每一个X值求它总体的均值的95%这一区间的时候，我们就可以得到一条？是不是把所有的点汇集起来，我们就可以得到关于总体回归函数的这样一个自信带了。同样我们对每个X求它个别值的95%的置信区间，就可以得到一个关于个别值的这样一个执行。在大家可以看到，我们这个图里面大家可以发现这个个别值它的自信贷总是要比总体均值。它的自信贷也就是说你的对于个体的预测值，它的这个带宽肯定是要大于对这个均值，或者说你对总体的均值，他的预测的自信带宽要更宽的。

发言人 01:25:03

那下面我们看什么。如果说我们用西塔来做这样一个各级预测的话，我们就可以把这些数据导入到我们这个模型里面。它包含几个变量，这是两个变量对吧？最简单的它存在着什么不同的地区作为样本。然后X是它的收入，Y是它的什么消费支出。

发言人 01:25:32

在state里面我们把数据导入进去之后就用这样一个很简单的这样一个代码说它有一个好处的，它这个代码词很多时候就是单那个单词的意思就是我regress。YX这样输进去，我们就构建的是什么Y等于贝塔0加贝塔1X加上六这样一个回归函数。从这样一个回归表里面，大家就可以看到，我会汇报出来我的这个样本的离差的情况。有总离差，还有什么我们相应的什么样本的什么观测值数量，the number of observation.

发言人 01:26:20

如果从这个什么表里面右上角来看的话，第一行number of of box就是我们这个样本数量有多少。但是他这里是取了31个地区的数据，所以它的样本观测值是31个，有个N是等于31。他会汇报什么？就是我们刚刚给大家讲的这个F检验的F值。那它的自由度是什么？1 29，那它算出来的F值会汇报给你那同样的F去检验的话，是不是把这个分布里面的概率计算，它也能换成P值进行检验，对吧？大家可以看到自己就有pro大于F大于临界值，他的这个概率，他也会把这个地址给你。

发言人 01:27:15

然后第四行他会给你汇报这个什么s school s point，就是我们这节课给大家讲的这什么这个礼盒优度，对吧？也就是说我们这个模型的线性，这样一个模型，它对样本观测点的代表程度或者拟合情况。同样他会把这个ADJX ground就adJusting，underground就调整后的这个礼盒优度他也会给大家汇报。大家就可以发现我调整后的你额度肯定是小于我的这个，对吧？那同样他也会汇报我们相应的这个标准差的情况。那在这个模型汇报的左边，他也会给大家汇报这个模型的做model的部分是它sox这一列展示的什么，我这个残差或者说我这个变差的来源工作，它就是我们这个模型解释的部分。然后下面就是我们这个残差的部分，所以它就包含两部分，一个是什么模型可以解释的部分，然后还有模型解释不了的残差的部分。

发言人 01:28:34

就我们刚刚那个图里面一个黄色的代表total总的它下面嘉禾的汇总的就是我们刚刚图里面那条黄色的线，吧？然后绿色的线呢就是表示在model里面。那同样的它每一个自由度，其他的标准物都给你汇报出来，你就开根号计算。然后最下面最主要的是我们这个模型的表现对吧？或者说我们就这个参数的估计的情况，他也给你输出出来了，他就是用我们这个什么计算的方式，OLS的方式给你估计出来，它估计出来X的系数是什么。我们找个同学来看一下。

发言人 01:29:29

国有盘。这个X后面就是我构建的这个模型是什么，或者模型的函数形式是什么呢？

发言人 01:29:49

他有借据下吗？yeah. 那我X后面的这个code是它的相当于是它的什么系数，对吧？这个相当于什么？0.61相当于是什么？系数你说我们估计穿这个贝塔吧？那这下面还有一个什么cos它是什么固定值？也就是说我们的截距项。所以2846.588，那就相当于是什么？我们这边的非常礼貌？估计出来的值大家可以看到，我们除了看系数的大小，还要看系数的是吗？

发言人 01:30:46

对，看完大小，看方向，对吧？正负，然后还有什么呢？

发言人 01:30:56

显著性水平对吧？我们这个显著性水平是多少呢？

发言人 01:31:09

159。57429显著性水平是有多大？是给定的这个阿尔法这里是不是我还没给对吧，就是我得去看我给多少，然后再让他看这个T值跟哪一个临界值进行比较，对吧？那假如说我给这个显著性水平是0.05的情况。0.57429是什么？

发言人 01:31:45

显著性水平。置信区间？那这个置信区间是不是就是我们之前给大家讲，它也是一种判别显著性的方式？那我现在假如是说算出来的这个制定区间，那你能看出来它是显著的还是不显说X对Y的影响，我在这个表里怎么看它是显著还是不显著的。三种方式，对吧？

发言人 01:32:22

对他不经过0？那就不显著对吧？好几个。我们看看有没有同学有其他的想法。邓宇。所以我说这个可能不是随机的那我们怎么判断说他汇报出来这个X的曲子，大家可以看到，第一个0.61是它的系数。第二个standard的error，它就是什么标准物，对吧？因为它的这个标准差，第三个是它的T值，第四个是这个P值。

发言人 01:33:00

然后第五行第五列和第六列就是我们置信区间。我怎么样来判断这个X跟Y它是不是显著的呢？就刚刚那位同学说的，除了我们看0.61的这个方向的大小之外，我们还要看X和Y之间的关系，对吧？

发言人 01:33:27

他俩显著吗？显著，那怎么你怎么看出来的呢？

发言人 01:33:45

对倒数第二行的这些汇报的结果，对吧？你具体从哪个词看出来的？

发言人 01:34:00

P值这个P值它是什么都是零了？也就是说我最后算出来这个码，我把实际的数据带到我们这个概率分布里面，结果算出来可以接受的这个部分的区域，它的面积都是很小，接近于零的。所以它是一个什么不可能事件？所以只能说是我们的原假设做了，原假设错了，也就意味着我们可以拒绝掉原假设说这个系数是显著？

发言人 01:34:34

那除了批次我们还可以看什么呢？是不是之前老师给大家讲过，大家把这个什么0.05、0.1、零点都给记下对吧？它就是几个数，1.68、2.58，对吧？那这个时候假使说我给的是0.01的显著性的话，显著性水显著水平的话，那我们这个T值是不是就跟2.58进行比较，那可以发现什么这个T值是吧？大大于多少？大于我的临界值，那我这个临界值2.58在这里吗？对吧？那它大于2.58也就意味着它落在了什么拒绝阈值的。所以我通过这个T值，如果大家把这个T临界值背下来了，你看T值是也可以看出来它是显著的那除此之外，你还可以从什么数值看出它是显著的呢？

发言人 01:35:44

置信区间对吧？那自信区间你怎么看待它是显著的呢？

发言人 01:36:08

是不是它都是一个正数，对吧？他他经过零吗？不经过，也就是说什么我95%的可能性我都娶不到你了，对吧？但是我就被他以95%的可能性给都取不到理由，也就意味着那你就取不到零了，那你就显著区别于零了，那你就是一个什么显著性的结果，对吧？

发言人 01:36:36

好，这个是好，请坐。这个是我们通常去看什么regress出来的结果，我们重点看一下这个X这一行的系数。除了除此之外我们可是不是可以看出来这个模型是不是显著的呢？或者说我想我看系数我们这些是不是还讲了，我们也看一下这个模型表现的好坏，对吧？那我们怎么看这个模型的好坏呢？韩国。

发言人 01:37:15

对你和优度看出来他这里是多少来着？那你可以解释一下这个0.9760吗？

发言人 01:37:33

不是说。

发言人 01:37:46

对，而且它是一个什么一次的一元的回归模型。所以相当于这百分之多少？97.6%的全部都有我这一个X来解释了对吧？那除了我们你和优度还有什么我们可以看出来这个模型表现的好坏呢？

发言人 01:38:23

调整后的R光跟R光是不是类似的，也是你和优度的方面。我们最后也讲了，这个还可以看是吗？F检验对吧？对，我们还可以看这个F简约。看这个F检验，你能看出来这个模型是显著的还是不显著呢？是显著的对吧？你怎么看的呢？

发言人 01:39:25

怎么看的呢？是不是跟T检验一样，就除了我们去对表那个F分布的表，他是不是也给出了什么crop大于F临界值的这个P值吧？那它的P值是也是零？那是不是我们可以不用去对比这个财富的这个是吗？临界值直接从这个什么第三就是右上角第三行这个prob大于F的这个概率值也可以看出来，就跟P检验一样，它这里的T值汇报出你比较临界值也可以通过P检验，那零可以看出来对吧？所以从上面的那个零是不是也可以看出来我们这个模型是什么显著的对吧？好几个，我们实验课就约在这周末，然后实验室是大家记一下是那个F118。

发言人 01:40:38

这是收签。

发言人 01:41:20

从早上第一节课开始，然后大家也可以自己带电脑，应该那里也有台式机。对。

发言人 01:42:23

摇动一下，这个好一点。

发言人 01:42:49

今天晚上。他笑笑。表。