统计计算功能在测量误差计算及公路检测中的应用

主讲人：王中伟（湖南交通职业技术学院）

【主要内容】

1．数组及电子表格的基本操作；

2．单变量统计计算（以测量误差计算和公路检测数据分析为例）；

3．双变量统计计算（线性回归分析）；

4．双变量统计计算（非线性回归分析）。

一、 统计计算

统计计算在工程测量中误差（精度）计算、工程检测数据处理中应用广泛。

1．统计计算需要样本数据

统计样本数据：$x_1, x_2, x_3, \cdots x_n$

【案例 1】对某段距离等精度独立丈量了 6 次，距离丈量值分别为 49.988m、49.975m、49.981m、49.978m、49.987m、49.984m，试计算其算术平均值 $\overline{x}$ 和一次丈量中误差 m 。

样本数据以怎样的形式输入和存储在计算器中？如何操作？

2．统计计算的主要参数

（1）求和：

$$\sum x = x_1 + x_2 + x_3 + \cdots x_n$$

（2）平均数：

$$\overline{x} = \frac{\sum x}{n} = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n}$$

（3）总体标准差、样本标准差（中误差）

总体标准差：$\sigma_n = \sqrt{\frac{\sum\left(x_i - \overline{x}\right)^2}{n}}$

样本标准差：$\sigma_{n-1} = \sqrt{\frac{\sum\left(x_i - \overline{x}\right)^2}{n-1}}$

标准差又称为均方差。

在测量误差理论中，经常用到测量值的中误差 m 这个概念，有这样两个公式：

$m = \sqrt{\frac{[\Delta\Delta]}{n}}$ 、$m = \sqrt{\frac{[VV]}{n-1}}$，前者是根据真误差的计算公式，后者是根据改正数的计算公式，后者的计算公式就是样本标准差。

（4）其它：中位数、最大值、最小值等

3．统计计算的主要内容

（1）统计参数计算

（2）线性回归分析

（3）非线性回归分析

二、 样本数据的输入、编辑和存储——数组及电子表格

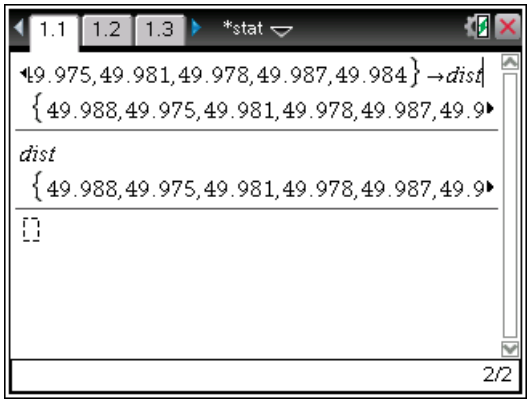
1．数组

数组是一系列数据的集合，用大括号括起来，各数据用逗号隔开。

【案例 1】对某段距离等精度独立丈量了 6 次，距离丈量值分别为 49.988m、49.975m、49.981m、49.978m、49.987m、49.984m，试计算其算术平均值 $\bar{x}$ 和一次丈量中误差 m 。

如：｛49.988,49.975,49.981,49.978,49.987,49.984｝

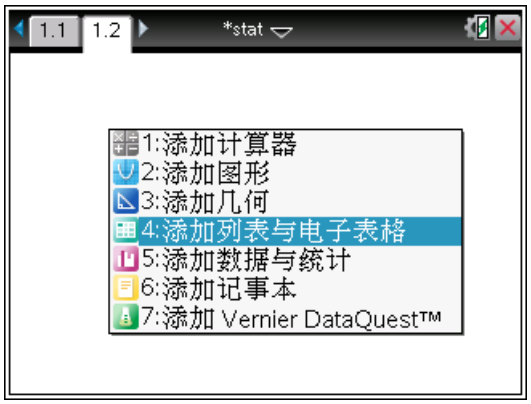
数组可以赋值给一个变量。如：｛49.988,49.975,49.981,49.978,49.987,49.984｝→dist

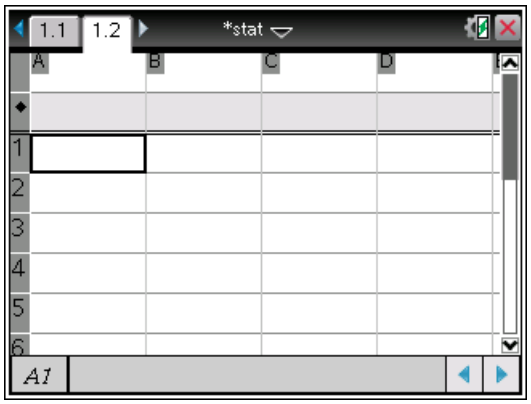


统计样本的数据使用数组方式进行输入和存储。

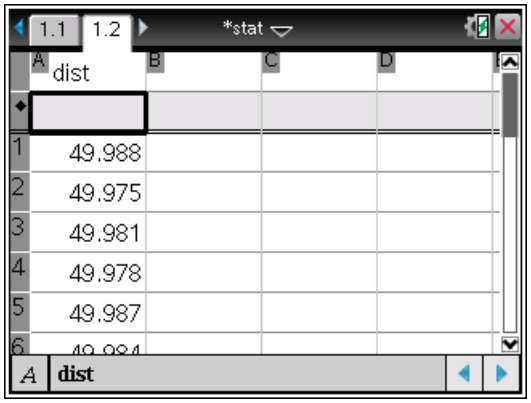
2．电子表格

电子表格是一个类似于 EXCEL 的表格视窗，按 ctrl+文档+ 4 键可添加一个或多个电子表格页面。

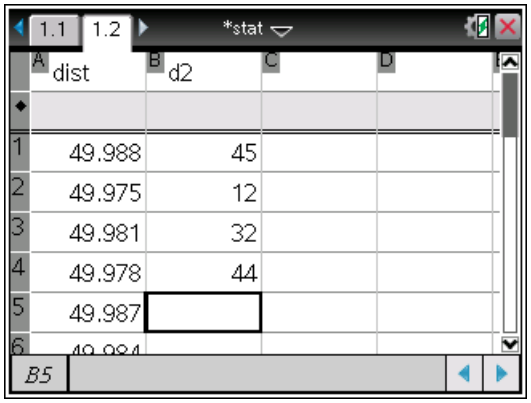


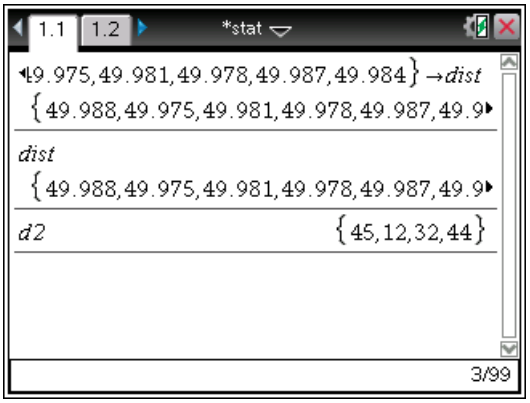


电子表格的一列，即可视为一个数组，两者是等同的。在电子表格中某列输入数组变量名称，即可在该列显示已经定义的数组数据。



同样，在电子表格中某列定义一个名称并输入数据后，也相当于定义了一个数组。





统计样本的数据，既可以通过数组定义来输入、修改和存储，也可以在电子表格中定义、修改和存储。显然，在电子表格中定义和修改比较直观。

三、 统计参数计算及其应用

【案例 1】对某段距离等精度独立丈量了 6 次，距离丈量值分别为 49.988m、49.975m、49.981m、49.978m、49.987m、49.984m，试计算其算术平均值 $\bar{x}$ 和一次丈量中误差 m 。

计算结果：$\bar{x} = 49.982\text{m}$，$m = 0.005\text{m}$

在将以上测量值赋值给数组变量 dist 后，可以使用以下三种方法计算出平均值和样本标准差（即测量中误差）。

1．利用菜单命令或者相应函数

菜单命令：

菜单—6:统计—3:数组运算—3:平均数

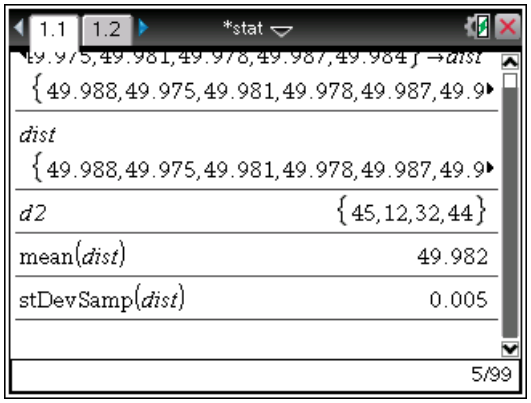
菜单—6:统计—3:数组运算—7:样本标准差

2．计算函数

使用菜单命令计算时会显示计算函数，若记得这些函数，也可以直接使用函数来计算：

平均值：mean(数组变量名)

样本标准差：stDevSamp(数组变量名)

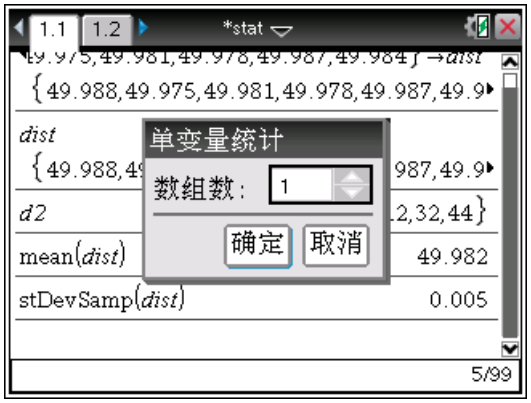


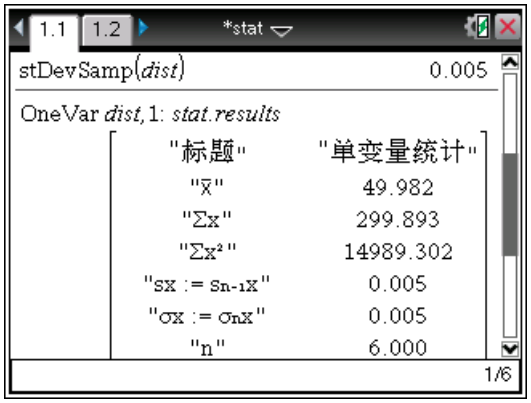
3．进行单变量统计功能计算全部参数

菜单命令：

菜单—6:统计—1:统计计算—1:单变量统计…

菜单文字后面带有“…”的表示会有对话框，依次按对话框提示操作即可。





【案例 2】某路段水泥混凝土路面板厚检测数据如表 1 所示，保证率为 95%，设计厚度 $h_d = 25\text{cm}$，代表值允许偏差 $\Delta h = 5\text{mm}$，试对该路段的板厚进行评价。

表1 水泥混凝土路面板厚度检测结果

序号	1	2	3	4	5	6	7	8	9	10
厚度 h_i(m)	25.1	24.8	25.1	24.6	24.7	25.4	25.2	25.3	24.7	24.9
序号	11	12	13	14	15	16	17	18	19	20
厚度 h_i(m)	24.9	24.8	25.3	25.3	25.2	25.0	25.1	24.8	25.0	25.1
序号	21	22	23	24	25	26	27	28	29	30
厚度 h_i(m)	24.7	24.9	25.0	25.4	25.2	25.1	25.0	25.0	25.5	25.4

本案例选自公路工程试验检测技术培训教材《路基路面试验检测技术》（张超等主编，人民交通出版社 2004 年 11 月出版）P77，例 3-31。

书中的解题过程如下，其中涉及到 t 分布概率系数需要查表。

例 3-31：某路段水泥混凝土路面板厚检测数据如表 3-10 所示。保证率为 95%，设计厚度 $h_d = 25\text{cm}$，代表值允许偏差 $\Delta h = 5\text{mm}$，试对该路段的板厚进行评价。

水泥混凝土路面板厚度检测结果　　表 3-10

序　号	1	2	3	4	5	6	7	8	9	10
厚度 h_i(cm)	25.1	24.8	25.1	24.6	24.7	25.4	25.2	25.3	24.7	24.9
序　号	11	12	13	14	15	16	17	18	19	20
厚度 h_i(cm)	24.9	24.8	25.3	25.3	25.2	25.0	25.1	24.8	25.0	25.1
序　号	21	22	23	24	25	26	27	28	29	30
厚度 h_i(cm)	24.7	24.9	25.0	25.4	25.2	25.1	25.0	25.0	25.5	25.4

解：经计算：　　$\overline{h} = 25.05\text{cm}$　　$S = 0.24\text{cm}$

根据 $n = 30$，$\alpha = 95\%$，查本书附录二(附表 2)得：

$$t_a/\sqrt{n} = 0.310$$

代表性厚度 h 为算术平均值的下置信界限，即：

$$h = \overline{h} - t_a/\sqrt{n} \cdot s = 25.05 - 0.310 \times 0.24 = 24.98\text{cm}$$

因为 $h > h_d - \Delta h = 24.5\text{cm}$，所以该路段的代表性厚度满足要求。

附录二　t 分布概率系数表

t 分布概率系数表　　附表 2

n	双边置信水平			单边置信水平		
	99%	95%	90%	99%	95%	90%
	$t_{0.995}/\sqrt{n}$	$t_{0.975}/\sqrt{n}$	$t_{0.95}/\sqrt{n}$	$t_{0.99}/\sqrt{n}$	$t_{0.95}/\sqrt{n}$	$t_{0.90}/\sqrt{n}$
2	45.012	8.985	4.465	22.501	4.465	2.176
3	5.730	2.484	1.686	4.201	1.686	1.089
4	2.921	1.591	1.177	2.270	1.177	0.819
5	2.059	1.242	0.953	1.676	0.953	0.686
6	1.646	1.049	0.823	1.374	0.823	0.603
7	1.401	0.925	0.734	1.188	0.734	0.544
21	0.621	0.455	0.376	0.552	0.376	0.289
22	0.604	0.443	0.367	0.537	0.367	0.282
23	0.588	0.432	0.358	0.523	0.358	0.275
24	0.573	0.422	0.350	0.510	0.350	0.269
25	0.559	0.413	0.342	0.498	0.342	0.264
26	0.547	0.404	0.335	0.487	0.335	0.258
27	0.535	0.396	0.328	0.477	0.328	0.253
28	0.524	0.388	0.322	0.467	0.322	0.248
29	0.513	0.380	0.316	0.458	0.316	0.244
30	0.503	0.373	0.310	0.449	0.310	0.239
40	0.428	0.320	0.266	0.383	0.266	0.206
50	0.380	0.284	0.237	0.340	0.237	0.184

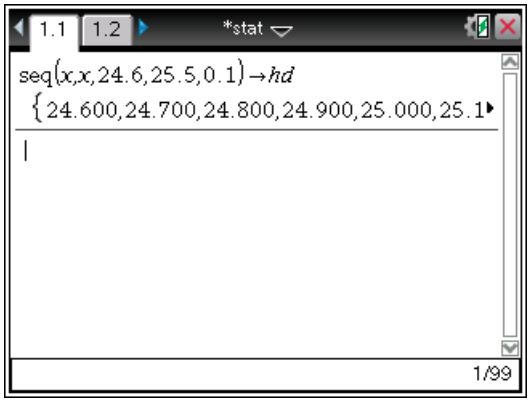
(1) 定义和输入统计数据：有两种方法，一是直接输入 30 个统计数据，二是根据表 1 整理统计数据如表 2 所示，建立统计数据和对应频数两个数组；

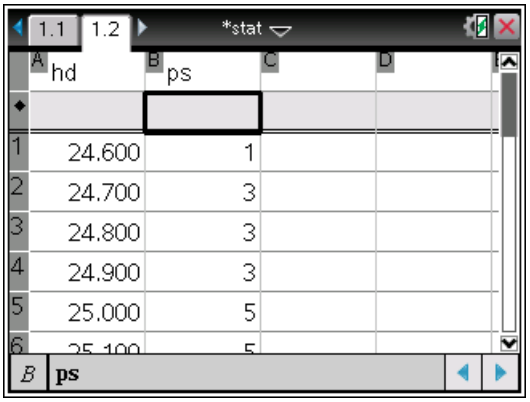
表2 【案例2】的统计数据和对应频数

序号	厚度 h_i(m)	频数
1	24.6	1
2	24.7	3
3	24.8	3
4	24.9	3
5	25.0	5
6	25.1	5
7	25.2	3
8	25.3	3
9	25.4	3
10	25.5	1

其中，我们发现统计数据很有规律，从 24.6m~25.5m，之间的步长是 0.1m，可以采用数列生成函数来生成。

函数格式是：seq(表达式，变量，下限值，上限值[，步长])



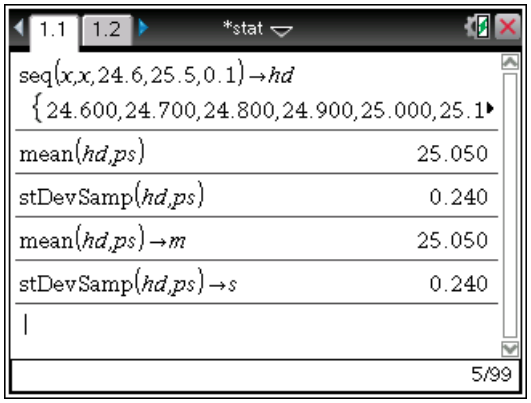


再在电子表格界面中显示 hd 数组，并建立一个对应的频数数组 ps 输入相应数据。

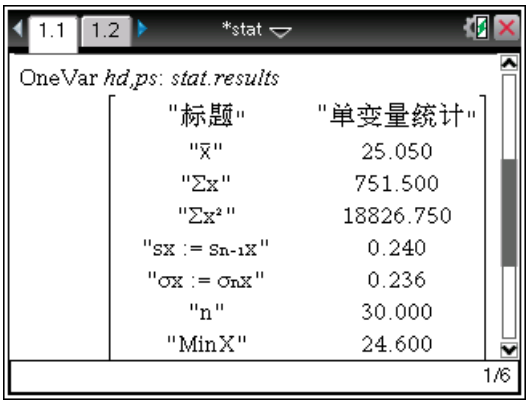
(2) 计算统计参数：平均值、样本标准差

计算结果为：$\bar{h} = 25.05\text{cm}$，$s = 0.24\text{cm}$。

直接利用函数进行计算：



或者使用单变量统计菜单命令计算：



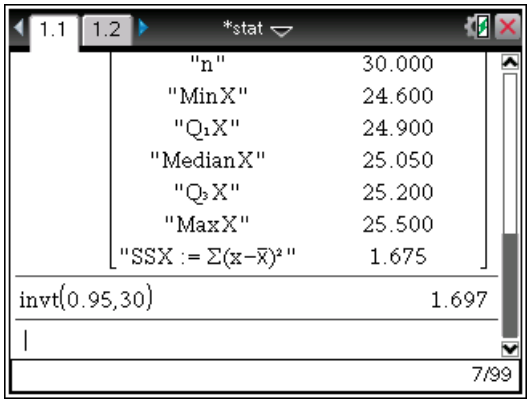
(3) 计算统计参数：t 分布概率系数

使用“反向 t”菜单命令或计算函数：

菜单命令：

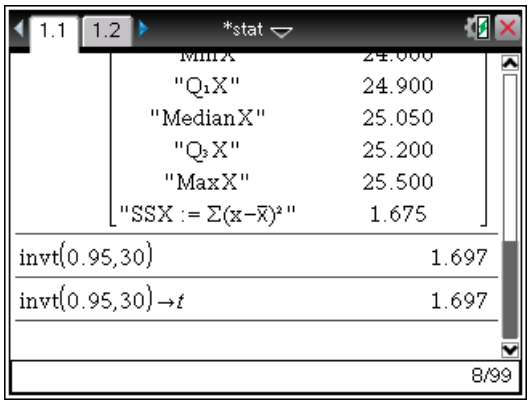
菜单—5:概率—5:分布—6:反向 t…

菜单—6:统计—5:分布—6:反向 t…



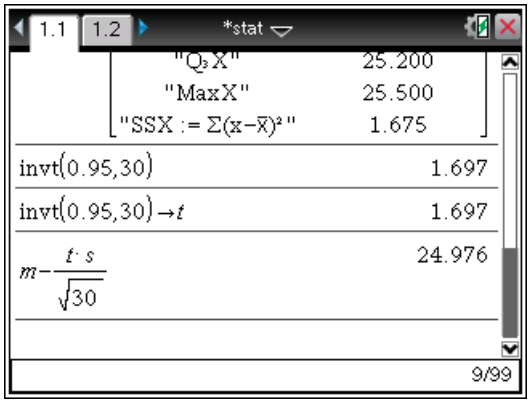
函数命令：

invt(面积或单边置信水平，自由度)



(4) 计算代表性厚度

$$h=\bar{h}-t_{0.95}/\sqrt{n}\cdot s=25.05-0.310\times 0.24=24.98\text{cm}$$



四、 双变量统计计算（线性回归分析）

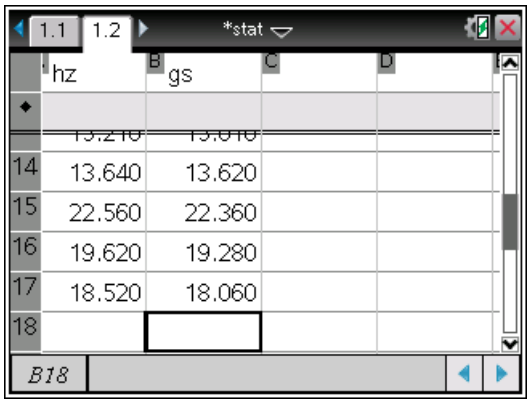
【案例 3】 某粘土路基施工中用核子密度仪法和灌砂法测量含水量的对比试验，试验结果如 3 所示，试对核子仪测定含水量进行标定。

表3 核子密度仪与灌砂法测定含水量对比试验结果

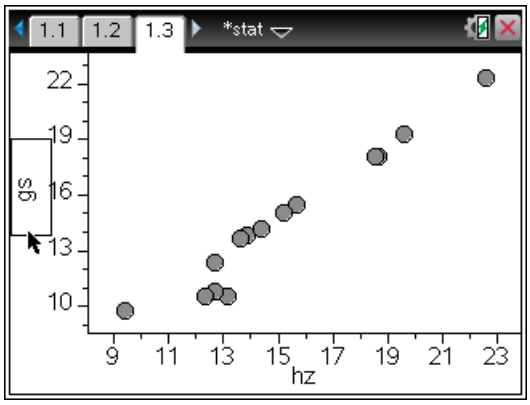
序号	核子密度仪	灌砂法	序号	核子密度仪	灌砂法
1	13.18	10.52	10	14.41	14.14
2	12.68	10.81	11	15.66	15.45
3	12.33	10.54	12	18.64	18.07
4	13.87	13.83	13	15.21	15.01
5	9.37	9.80	14	13.64	13.62
6	13.87	13.80	15	22.56	22.36
7	15.21	15.05	16	19.62	19.28
8	13.87	13.83	17	18.52	18.06
9	12.66	12.33			

背景知识：核子密度仪（又称核子密度湿度仪，简称核子仪）一般用于测定路基或路面材料的密度和含水量，并以此计算施工压实度，是一种施工质量的现场快速评定方法，但该法测定的结果与标准方法（灌砂法）测定的结果有一定的差异，因此使用核子密度仪测定的结果一般要与灌砂法测定的结果进行标定，建立两种方法之间的相关关系（线性相关），建立回归方程。用于标定的数据至少应有 15 组，其相关系数 r 应不小于 0.9。

(1) 定义和输入统计数据：分别建立 hz、gs 两个数组分别输入核子密度仪和灌砂法两组数据；



(2) 添加数据与统计页面，并定义（或修改）横轴为 hz（核子密度仪），竖轴为 gs（灌砂法），生成散点图：



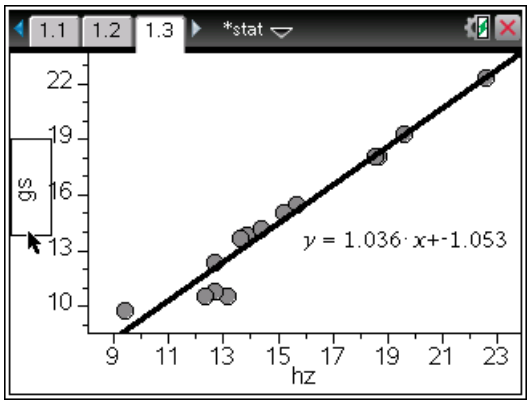
从生成的散点图可以看出两者呈现比较明显的线性关系。

(3) 回归分析计算：

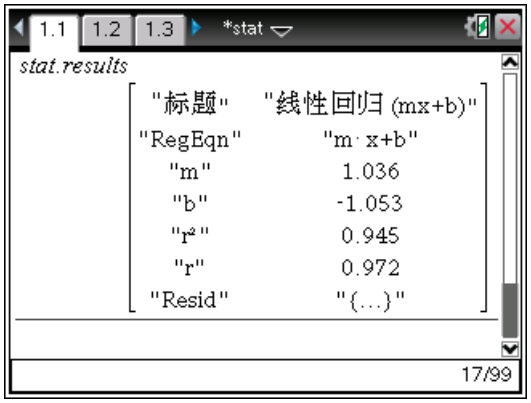
按菜单命令：

菜单—4:分析—6:回归—1:显示线性回归(mx+b)

图形显示回归线条和回归公式，即：$y_{砂} = 1.036x_{核} - 1.053$



输入 stat.results 可查看回归分析的所有结果：



线性回归的相关系数 $r = 0.97 > 0.9$，呈明显的线性相关。

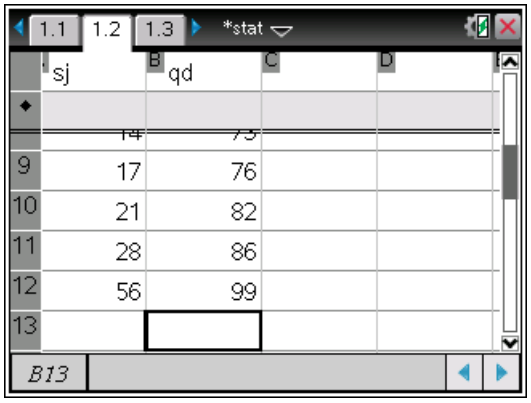
五、 双变量统计计算（非线性回归分析）

【案例 4】混凝土的抗压强度 R 随养护时间 t 的延长而增加，表 4 列出了一批混凝土的养护时间 t（天）与抗压强度 R（kg/cm^2）的测试数据，试根据这些数据分析两者之间的关系，求出回归方程式。

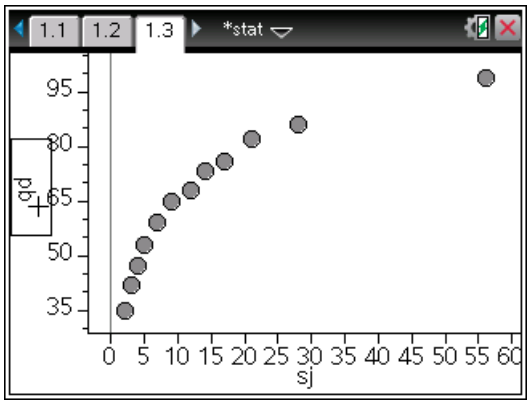
表4 混凝土的养护时间与抗压强度的测试数据

序 号	1	2	3	4	5	6	7	8	9	10	11	12
时间 t(d)	2	3	4	5	7	9	12	14	17	21	28	56
抗压强度 R(kg/m^2)	35	42	47	53	59	65	68	73	76	82	86	99

(1) 定义和输入统计数据：分别建立 sj、qd 两个数组分别输入时间和抗压强度两组数据；

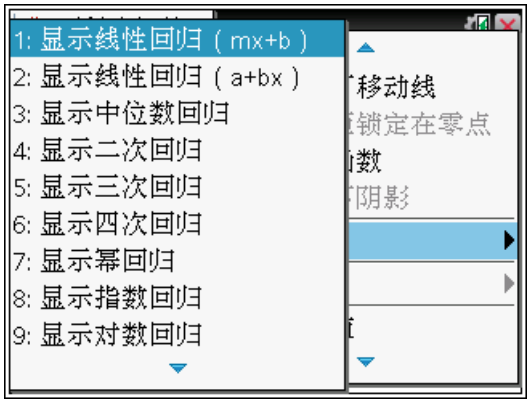


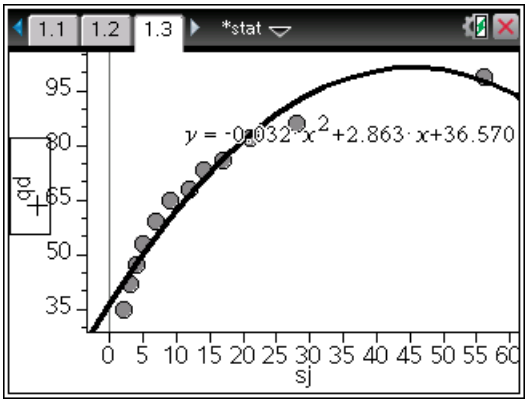
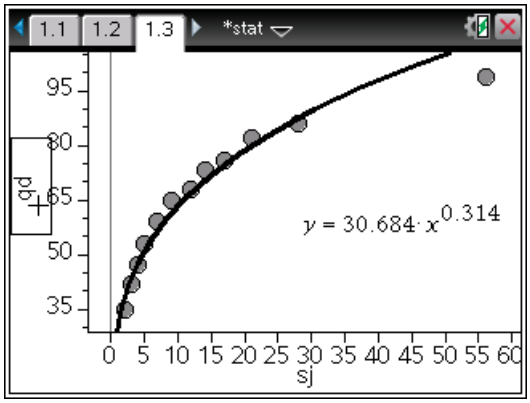
(2) 定义（或修改）横轴为 sj（时间），竖轴为 qd（抗压强度），生成散点图：

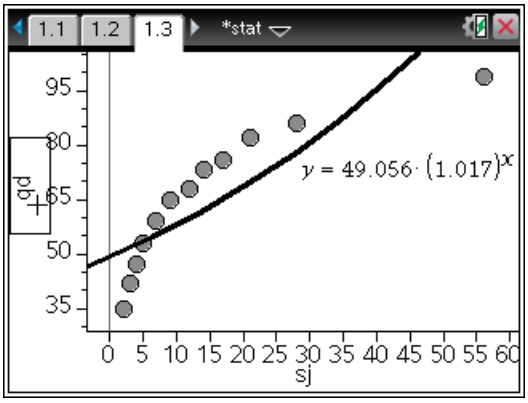
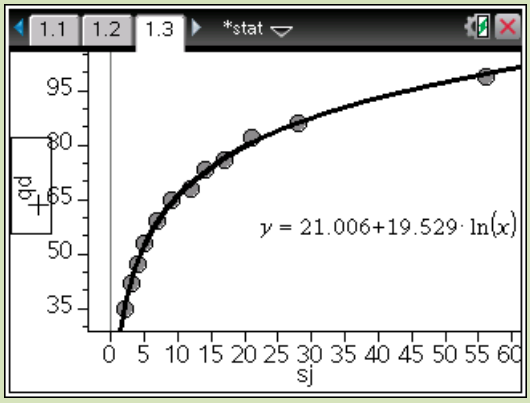


很显然，散点图不呈线性关系，而呈非线性关系，但具体是哪种曲线呢？

(3) 判别或尝试各种曲线回归类型：

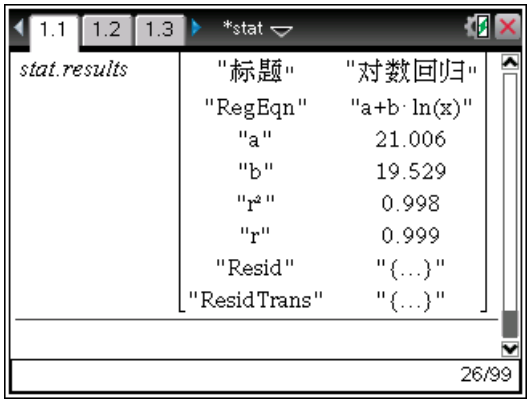


二次回归	幂回归
	

指数回归	对数回归
	

观察各种图形，可确定，本例中混凝土的养护时间 t 与抗压强度 R 之间呈明显的对数关系，回归方程式为：$R = 21.006 + 19.529\ln(t)$。

输入 stat.results 可查看回归分析的所有结果：



相关系数 $r = 0.999$，呈高度相关特性。

课后练习题：

【练习 1】用 S3 水准仪观测 A、B 两点间的高差 h_{AB} 共 10 次，分别为 2.749、2.753、2.750、2.748、2.752、2.747、2.751、2.750、2.749、2.751（单位：m），试计算 h_{AB} 的算术平均值、观测值中误差和算术平均值中误差。

【练习 2】某公路桥梁，桥位上游附近，有一个水文站（乙站，分析站），搜集到 1955 年至 1982 年的年最大洪水流量资料，但其中 6 年，因故缺少资料，只有 22 年的实测流量资料。在临近流域的河流上，也有一个水文站（甲站，参考站），可以搜集到 1951 年至 1982 年连续 32 年的年最大洪水流量资料。两流域的特征基本相似，拟利用甲站的流量资料对乙站的流量资料进行插补和延长。

试进行相关分析，计算回归方程式，并对乙站的流量资料进行插补和延长。

序号	年份	甲站流量 Qx(m^3/s)	乙站流量 Qy(m^3/s)		序号	年份	甲站流量 Qx(m^3/s)	乙站流量 Qy(m^3/s)
1	1951	964	—		17	1967	4440	—
2	1952	2299	—		18	1968	2702	—
3	1953	1645	—		19	1969	755	600
4	1954	1938	—		20	1970	2000	1530
5	1955	2643	2000		21	1971	2420	2170
6	1956	2757	2380		22	1972	1870	1650
7	1957	2549	2100		23	1973	1060	—
8	1958	3359	2600		24	1974	3710	—
9	1959	3997	2950		25	1975	1900	1300
10	1960	4093	—		26	1976	2180	1850
11	1961	2840	2500		27	1977	1350	900
12	1962	1160	1000		28	1978	4920	—
13	1963	1560	1100		29	1979	2385	1900
14	1964	2470	1360		30	1980	1200	1080
15	1965	2150	1480		31	1981	1300	1010
16	1966	3025	2250		32	1982	2360	1700

【练习 3】某新建公路的路基施工中，对其中某一路段进行压实质量进行检查，压实度检测结果分别为 96.4%、95.4%、93.5%、97.3%、96.3%、95.8%、95.9%、96.7%、95.3%、

95.6%、97.6%、95.8%、96.8%、95.7%、96.1%、96.3%、95.1%、95.5%、97.0%、95.1%，请按保证率 95%计算该路段的代表性压实度，并判断该路段的压实质量是否符合要求（压实度标准为K_0=95%）。