

GRAPHER 在统计分析中的应用

◆ 和向丽

GRAPHER 是久负盛名的 SURFER 的姊妹软件, 为 GOLDEN 公司的产品, 目前最新版本为 5.04 版。该软件的主要特点: 一张图内可画无数条线, 每条线可达 32000 点; 可用线性坐标、对数坐标; 线型符号颜色都可定义; 支持剪贴板功能; 具有 Curve Fitting 功能, 文字可用上下标数学符号..... 在绘图时更加得心应手。除此之外, GRAPHER 还有以下特点:

1. 具有完善的数据输入接口。在 Worksheet 中可以读入几乎所有常用的数据文件, 包括 MICROSOFT EXCEL 文件 (*.XLS)、数据文件 (*.DAT)、纯文本文件 (*.TXT) 文件等。

2. 具有方便的数据计算功能。利用 Worksheet 的 Data-Transform... 能够方便地进行各种计算, 包括各种函数计算。

3. 具有简单易用的统计功能。Worksheet 的统计功能 (Statistics) 可以对所选择的任意段数据进行统计分析, 包括“数据个数、平均值、标准偏差、最大值、最小值、变异系数、峰度系数、偏斜系数”等, 并能将统计的结果利用剪贴板粘贴到任何合适的位置。

4. 具有强大的统计制图功能。GRAPHER 能够做出各种 2 维、3 维的折线图、散点图、饼图、玫瑰花图等统计图。

5. 具有完备的数据输出接口。GRAPHER 支持十余种输出文件格式, 除了常用的图片文件格式 (*.BMP、*.JPEG、*.TIF) 外, 还支持 *.DXF 文件格式, 从而使 GRAPHER4 中的图形能够是目前广泛使用的矢量化软件 MapGIS、AutoCAD 等软件调用。

GRAPHER 在工程技术领域应用非常广泛, 被誉为二维科学绘图软件的主流。在工资统计中往往是 EXCEL 独霸天下, 其实在某些方面 GRAPHER 比 EXCEL 使用更为方便, 灵活应用能够起到事半功倍的效果。本文用 GRAPHER4 对单位 2005 年 4 月 100 名职工工资 (图表略) 进行了统计分析。

统计工资参数

打开“工资.XLS”, 点“Data-Transform...”, 选择要计算的参数, 点“OK”, 计算参数见表 2。

由工资参数图表 (附后) 可见, GRAPHER 统计了常用的样品数、平均值、最大值、最小值等各项参数, 并给出了数据分布正态分布参数, 从这些参数可以对工资情况有个大致了解: 平均工资 985.30, 高于所在城市 800 元的平均工资水平。平均值与中位数的联合使用可用于了解数据的频数分布形态, 如果均数与中位数相等, 观察值通常呈对称分布; 如果均数大于中位数, 观察值分布偏向右侧, 为正偏态分布; 如果均数小于中位数, 观察值分布偏向左侧, 为负偏态分布。表 2 中平均工资大于中位数, 表明工资数据呈正偏态分布型式。变异系数 (Coefficient of variation) 主要反映数据的分异程度, 为无量纲数据, 便于资料间的相互比较, 它常用于比较均数相差悬殊的几组资料的变异程度, 以及比较间度量衡单位不同的几组资料的变异程度。从表中可见工资的变异系数仅为 0.19, 表明工资是基本“均匀”的, 换言之, 也就是说工资的差异并不是太大 (变异系数分级: >2.0 极强分异, 1.00—2.00 强分异, 0.75—1.00 分异, 0.50—0.75 弱分异, 0.25—0.50 不均匀, <0.25 均匀)。

河南省创佳公司 2005 年 4 月份工资参数统计结果

Number of values(样品数)	100
Sum(和)	98530
Minimum(最小值)	600
Maximum(最大值)	1600
Mean(平均值)	985.3
Median(中位数)	925
First quartile(第1四分位数)	840
Third quartile(第3四分位数)	1132.5
Standard error(均值标准误差)	19.59963
95% confidence interval(均值95%置信区间)	38.89155
Variance(方差)	38414.56
Standard deviation(标准误差)	195.9963
Coefficient of variation(变异系数)	0.19892
Skew(偏度)	1.152
Kurtosis(峰度)	1.208
Kolmogorov-Smirnov stat(正态分布Kolmogorov-Smirnov吻合度)	0.231
Critical K-S stat, alpha=0.05(95%显著水平的K-S统计临界值)	0.134

生成各种统计图

1 步长图(Step Plot)

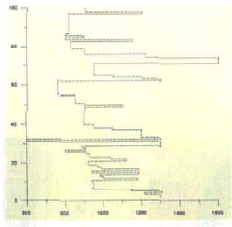


图1 直观反映数据的分布情况

2 饼图(Pie Chart)

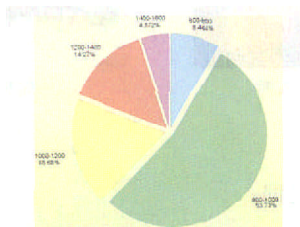


图2 反映不同工资所占的比例

3 三维柱状图(3D Bar Chart)

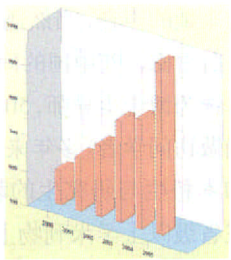


图3 为反映不同年度平均工资增长幅度

4 三维带状图(3D Ribbon)

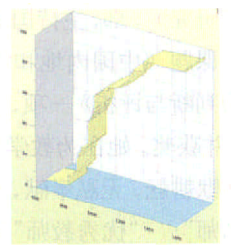


图4 反映不同工资额的累计频率

5 洛伦茨(Lorenz) 曲线图

洛伦茨曲线图最早由美国洛伦茨博士(Dr. M.O.Lorenz) 提出,其实质为累计频数(或频率) 分布曲线,可用以研究财富、土地和工资收入的分配是否公平,表1数据的洛伦茨曲线图如图4,图中AC线称为绝对平等线,折线ABC称为绝对不平等线。由图可见,实际工资线与绝对平等线几乎一致,表明工资分配高度平均,这种高度平均的工资分配模式对于充分调

动员工的工作积极性是不利的。

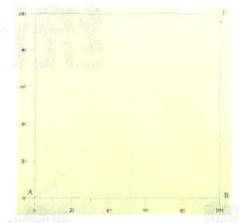


图5 工资数据洛伦茨曲线图

工资分布型式分析

根据数理统计理论,对于一个随机变量 X ,只要它的各次观测值之间的判别是由无数因素引起的,而这些因素中没有一个是突出的,则它的概率分布自然地趋向正态分布[5]。对数据做频率分布直方图能够大致了解其分布型式。图5为表1数据的频率分布直方图,从图中可见,工资数据基本呈略为正偏(右偏)的正态分布,工资数据主要分布在800—1000元(组中值为900),占全部工资分布的60%,工资在600元以下(含600元)约占总数的1%,1000元以上工资约占30%,工资总体分布比较合理。

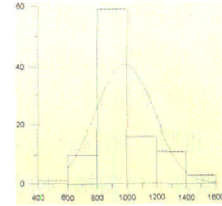


图6 工资数据频率分布直方图

统计图形输出

在GRAPHIER中生成的各种统计图形,可以导出(Export)为JPG、TIF、BMP、DXF等20种文件格式。其中最常用的为图片文件格式(*.JPG)和AutoCAD DXF文件格式(*.DXF),前者可以在Word文档和PowerPoint文档中引用,后者可以被目前普遍使用的矢量化软件MapGIS、AutoCAD、CorelDRAW调用。

点“File”-“Export...”,选择导出的文件格式,用于输出的图片文件格式(*.JPG)水平(Horizontal)分辨率和垂直(Vertical)分辨率设为150DPI即可,AutoCAD DXF文件格式(*.DXF)选默认设置。

结论

1. GRAPHIER不仅可用于工程技术领域,在工资数据的统计分析中也可以得于广泛应用。

2. 分析GRAPHIER的各种统计参数及统计图形,可以用于了解工资数据的分布情况,为制定工资政策提供决策依据。■

【作者简介】和向丽,女,河南西平人。1993年毕业于福建地质学校,2000年毕业于河南财经学院,现在河南省地质调查院从事财务工作,会计师。