A Plot Method for "htest" Objects

Richard M. Heiberger    G. Jay Kerns
Temple University    Youngstown State University

The numerical results of many statistical tests in R are stored in an "htest" object. The print method for the class displays a table. We have written a generic plot.htest function for the class that calls the graphing functions in the HH package.

> z <- rnorm(20, 2, 3)
> t.test(z)

One Sample t-test

data:  z
t = 2.4213, df = 19, p-value = 0.02564
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  0.1986526 2.7321318
sample estimates:
  mean of x
  1.465392
> z <- rnorm(20, 2, 3)
> t.test(z)

    One Sample t-test

data:  z
    t = 2.4213, df = 19, p-value = 0.02564
alternative hypothesis:
    true mean is not equal to 0
95 percent confidence interval:
    0.1986526  2.7321318
sample estimates:
    mean of x
    1.465392

> plot(t.test(z))
>
> vt <- var.test(x, y, alt="greater")
> vt

F test to compare two variances
data:  x and y
F = 3.6, num df = 6, denom df = 18,
p-value = 0.01598
alternative hypothesis:
  true ratio of variances
  is greater than 1
95 percent confidence interval:
  1.35272   Inf
sample estimates:
  ratio of variances
    3.6

> plot(vt)
We can specify the graph from the Rcmdr menu.
RExcel

Microsoft Excel is the most widely used spreadsheet program. Many of our clients and students use it as their data management system and as their working environment.

On Windows RExcel and statconnDCOM (http://rcom.univie.ac.at) access COM—the Microsoft interprocess communications system, and seamlessly integrates the entire set of R’s statistical and graphical methods into Excel.

Therefore these "htest" graphing methods are available through RExcel.

The normal example is from the recent book R through Excel by Richard Heiberger and Erich Neuwirth. It is possible to build half the introductory course on this one graph.

There are many more examples included with the software. There are several other talks at this UserR! 2010 conference illustrating applications of RExcel.
Placing values in the cells in Excel provides live control of the curve displayed in the R graph. The slider on $\mu_1$ in Excel, smoothly moves the normal curve centered at $\mu_1$ and adjusts the corresponding area illustrating $\beta$, the probability of the Type II Error.
Placing values in the cells in Excel provides live control of the curve displayed in the R graph. The slider on $\mu_1$ in Excel, smoothly moves the normal curve centered at $\mu_1$ and adjusts the corresponding area illustrating $\beta$, the probability of the Type II Error.
A Plot Method for "htest" Objects

July 21, 2010
Richard M. Heiberger and G. Jay Kerns

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Distribution</td>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Input</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>observed F</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( \nu_1 = ) numerator degrees</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>( \nu_2 = ) denominator degree</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>( \alpha = ) Prob(Type I error)</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>rejection side</td>
<td>right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ncp = noncentrality</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>( \theta = ) multiple of central F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>nn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>( s^2_x )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>( s^2_y )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>( R_0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Highlight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>Type II Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>Power</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Output</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>critical value—left</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>critical value—right</td>
<td>2.661</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>p-value</td>
<td>0.0160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>( \beta = ) Prob(Type II error)</td>
<td>0.3517</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>power=1-( \beta )</td>
<td>0.6483</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Done: distrPlot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>Graph on Top</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>data scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>labelbox</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A Plot Method for "htest" Objects

July 21, 2010
Richard M. Heiberger and G. Jay Kerns

F density, \(\nu_1 = 6, \nu_2 = 18\)
Non–centrality parameter \(\lambda = 14\)
A Plot Method for "htest" Objects

July 21, 2010
Richard M. Heiberger and G. Jay Kerns

F density, $\nu_1 = 6, \nu_2 = 18$
Non–centrality parameter $\lambda = 14$

$F_{\text{crit.}} = 2.661$
$F_{\text{obs.}} = 3.6$
$\alpha = 0.050$
$\beta = 0.352$
$1 - \beta = 0.648$
$p = 0.016$
F density, $\nu_1 = 6$, $\nu_2 = 18$

Non-centrality parameter $\lambda = 19$

$\alpha = 0.050$
$\beta = 0.200$
$1 - \beta = 0.800$
$p = 0.016$
F density, $\nu_1 = 6$, $\nu_2 = 18$

Non-centrality parameter $\lambda = 14$
F density, $\nu_1 = 6$, $\nu_2 = 18$

Non-centrality parameter $\lambda = 19$

$s_x^2 = 7.2$, $s_y^2 = 1$

$\beta = .200$, $1 - \beta = .800$

$\alpha = .050$, $p = .016$
F density, $\nu_1 = 6$, $\nu_2 = 18$

Non-centrality parameter $\lambda = 19$
Graphical Design

1. Color choice
2. Outline of $p$-value area
3. Alternate axes

In the normal and $t$ plot, we show the $\bar{x}$-scale, the $z$-scale under the null, and the $z_1$-scale under the alternative. We can show the data scale for the $F$ and $\chi^2$. 
Design Questions

I invite discussion afterwards on these topics.

1. Color scheme
2. Legend vs Labels
3. Static vs Dynamic
4. Paper vs Screen
Conclusions

Dynamic graphs of hypothesis tests are an excellent way to understand the material and to teach the material.

They can be used in production as part of experimental design. Inspecting these graphs can help in the determination of sample size.
References

You can download this paper for no charge if your library subscribes.

http://www.springer.com/978-1-4419-0051-7