Understanding P-values Through Simulation

This applet (programmed by German Molina) simulates the proportion of times that the null hypothesis is true when the p-value is in the range `from` □ `to` □. If one wishes to find the proportion of times the null hypothesis is true for a given p-value (e.g., $p = 0.05$), one should choose a small range (e.g., from 0.049 to 0.05).

The applet considers testing of $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$, where μ is the mean of a normal distribution having standard deviation one. (Any other standard deviation would yield equivalent results.) Normal observations of `sample size` □ are taken, and testing is based on the usual z-statistic $\sqrt{n}\,|\overline{X}|$, where n is the specified sample size and $\overline{X}$ is the sample mean. The simulation creates a long series of such tests, and simply counts how often H_0 is true and false, whenever the p-value is in the specified range. These counts are given in the boxes on the right of the graph and are also represented in the graph, which can be chosen to be either of columnar or line `graph type`.

Pushing the button `Add i` will cause the program to run until `i` tests have occurred with p-values in the desired range. Any of the `Add i` buttons can be pushed again, and the results will be added to the existing totals. Pushing the `Continuous` button will cause the program to run continuously. One can switch between the `Add i` buttons and the `Continuous` button, and totals will stay accumulated.

To start over, one pushes `End/Refresh`. A summary of the last run is then presented on the screen. If several different runs are done, the results are saved and can be accessed by the `Results` button.

To run the applet, one must also choose

- the `% of true nulls` that are to be generated, i.e., the proportion of null hypotheses, H_0, that are initially chosen to be true in the sequence of simulated tests;
- the values of the normal means μ that arise under the alternative hypotheses, H_1, in the sequence of simulated tests.

If one desires to choose all the alternative means to be at a fixed point, choose `point mass` under `Distribution of alternatives` and then choose the desired location of the alternative means using `with location` □.

The alternative means can also be randomly generated from any of the distributions under `Distribution of alternatives`. These distributions are all given in standardized form. Thus selecting `normal` will result in the standard normal distribution. One can shift this distribution to a different location by entering the desired location using `with location` □, and can change the scale of this distribution using `and scale` □. Thus entering a `location` of 2 and a `scale` of 3 will result in the normal distribution with mean 2 and standard deviation 3. One can specify the location and standard deviation either in real units, or in `sample standard deviation units` (i.e., in units of $1/\sqrt{n}$), by checking the appropriate box. Finally, the `See Means` button shows the values of the generated alternative means μ that happened to have p-values in the indicated range.

When $p \approx 0.05$, it is interesting to note that the final percentage of true nulls will usually exceed the initial percentage, unless the values of μ under the alternatives are chosen very carefully. Indeed, finding the choices of alternatives that minimize the final percentage of true nulls is an interesting exercise. As an example of what one will find if the initial percentage of true nulls is 50%:

- the final percentage of true nulls is *at least* 22% when the p-value is between 0.049 and 0.050;
- the final percentage of true nulls is *at least* 6.5% when the p-value is between 0.009 and 0.01.

Final Note: Essentially the same results can be seen to hold when the point null hypothesis, $H_0 : \mu = 0$, is replaced by a small interval null hypothesis, $H_0 : -\varepsilon < \mu < \varepsilon$.