Disclaimer:
Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.
Statisitical Testing for MCNP

Anthony J Zukaitis XCP-3
Art Forster XCP-3
Rick Picard CC6-6

July 13, 2021
Background - Ideology

- Statistical testing is used when exact answers cannot be replicated.
- The sampling distribution of the mean of a variable will be normally distributed.
- Compare tallies from two random number sequences (rand card).
- Use a specific tally as ensemble via bins.
Testing

Tally bins should be well converged

- Valid Confidence Intervals
- Regular Tallies: Std. err $\leq 0.05$
- Point detectors: Std. err $\leq 0.1$
- VOV $< 0.1$ recommended but not always available.
StatsTools

Ideology encapsulated into two Python tools

- Python Dependencies: MCNPTools, scipy
- do_mctal_stats.py
- do_meshtal_stats.py
do_mctal_stats.py

bash-4.2$ do_mctal_stats.py --help
usage: do_mctal_stats.py [-h] --new NEW --ref REF
      --tally TALLY [--err ERR] [--pcrit PCRIT] [--report]

optional arguments:
  -h, --help          show this help message and exit
  --new NEW           New mctal file
  --ref REF           Reference mctal file
  --tally TALLY       Tally number to process
  --err ERR           Allow values equal or less than this relative error
                      (default: 0.05)
  --pcrit PCRIT       Critical pvalue to use for statistical tests
                      (default: 0.005)
  --report            Report bin by bin analysis
Statistical Significance

- Statistically probable measurements
- Validation "Good"="Not Bad"
- Critical levels for probability 0.05, 0.01, 0.005
Statisical Tests

- $Z^2, \log P$ summed (both follow $\chi^2$ distributions)
- $Z, P$ binned (Std. normal vs uniform)
- $Z, P$ Kolmogorov Smirnov (erf vs linear)
Z-Values

The Z-value statistic of normally distributed values follow a standard normal distribution.

- \[ Z_i = \frac{x_i - y_i}{\sqrt{s_{x_i}^2 + s_{y_i}^2}} \]
- \( x_i \) is the tally result obtained via one random number sequence
- \( y_i \) is the tally result obtained via a different sequence
- \( s_{x_i,y_i} \) is the standard deviation
- \( Z_i \approx 0 \) values are thrown away (not statistically different)

The standard deviation is computed from the MCNP estimated standard error
P-Values

- $P$-value defines the probability of a measured $Z$ or higher
- $P = \frac{1}{2} \left(1 - \text{erf} \left( \frac{|Z|}{2} \right) \right)$
P-Values

- $P$-value defines the probability of a measured $Z$ or higher
- $P = \frac{1}{2}(1 - \text{erf}(\frac{|Z|}{2}))$
Chi squared background

- $Z_i$ is normally distributed
- $\chi^2 = \sum_{i=1}^{N} Z_i^2$
- $dof = NN =$ number of bins
- Then $\chi^2$ follows a chi-square distribution
Chi squared background

- p-values define the probability of a measured $\chi^2$ or higher
Chi squared background

- \( p \)-values define the probability of a measured \( \chi^2 \) or higher
Chi squared summed testing

- \( \chi^2 = \sum_i^N Z_i^2 \)
- \( \chi^2 = -2 \sum_i^N \log (2 \times P_i) \)
- From each computed and \( \chi^2 \) a p-value can be computed with respect to the \( \chi^2 \) distribution
Chi squared binned testing

- specify a number of bins (nbins=21,20) for $Z, P$
- prepare histogram bins for $Z, P$
- Compute bin boundaries such that the volumes of the bins are equal.
Chi squared binned testing

- $O_j$ is the number of $Z,P$ within bin $j$
- $\chi_j = \frac{(O_j - E_{exp})^2}{E_{exp}}$
- $\chi^2 = \sum_{j}^{nbins} \chi_j$
- A p-value can be computed for this $\chi^2$
Kolmogorov-Smirnov $Z, P$ values testing

- Construct CDF of sorted $Z$, $P$ values
- Construct corresponding CDF
  - CDF for $Z$ is $\frac{1}{2} \text{erf}(\frac{Z}{\sqrt{2}})$
  - CDF for $P$ is $P$ (linear).
- Look for largest separation between the two
- This distance is the d-value
- Compute corresponding p-value
Test Problem bas-01

SDEF POS=0 0 0 RAD=D1 ERG=D2 TME=D3
  
SI1 0 1 $ RAD distribution
SP1 -21 2
  
SI2 S 21 22 $ ERG distribution
SP2 1 2.
SP21 -3 $ Watt spectrum
SI22 L 1.1 1.5 2.0 $ Discrete lines
SP22 2 1 3.
  
SI3 H 10 100 $ Pulse
SP3 0 1
  
F1:N 1 $ ERG and TME tallies
E1 1E-2 63i log 10
T1 5 63 i 110
  
F11:N 1 $ ERG Only
E11 1E-2 255 i log 20
Test Problem bas-01 - Surface Tally 1

- Time - Uniform pulse distribution
- Energy - 3 lines + Watt Spectrum
- Space - Spherical Volume $r = 1\text{ cm}$
Test Problem bas-01 - Surface Tally 11

- Energy - 3 lines + Watt Spectrum

![Graph showing energy spectra](image)
## Test Problem bas-01

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing tally</td>
<td>11</td>
</tr>
<tr>
<td>Nonzero scores</td>
<td>256</td>
</tr>
<tr>
<td>Critical Z</td>
<td>2.6600674686174592</td>
</tr>
<tr>
<td>Warning! Large Z</td>
<td>2.815806122121412</td>
</tr>
<tr>
<td>Warning! Large Z</td>
<td>2.829843781428351</td>
</tr>
<tr>
<td>Val ref</td>
<td>0.0029065</td>
</tr>
<tr>
<td>Val new</td>
<td>0.002975</td>
</tr>
<tr>
<td>Err ref</td>
<td>0.0059</td>
</tr>
<tr>
<td>Err new</td>
<td>0.0058</td>
</tr>
<tr>
<td>Total scores</td>
<td>257</td>
</tr>
<tr>
<td>Usable scores</td>
<td>196</td>
</tr>
<tr>
<td>Unusable scores</td>
<td>60</td>
</tr>
<tr>
<td>Unusable matching zero scores</td>
<td>1</td>
</tr>
<tr>
<td>Unusable matching nonzero scores</td>
<td>0</td>
</tr>
</tbody>
</table>
Test Problem bas-01

Ref scores with errors <= 0.05 : 196
New scores with errors <= 0.05 : 197
Errors outside crit ratio : 3.0 count: 0
Minimum non zero error over all ref data: 0.0004
Minimum non zero error over all new data: 0.0004
zspace_binned: scores: 196 chisq: 22.73479444149072 dof: 20 pvalue: 0.30200355809947327
z_summed_test: chisq: 225.5458350589523 dof: 196 pvalue: 0.07261966846939263
zspace_ks: scores: 196 D: 0.11914784582184068 pvalue: 0.006993056520287106
pspace_binned: scores: 196 chisq: 15.836734693877549 dof: 19 pvalue: 0.6681419596523805
pspace_summed: scores: 196 chisq: 440.42402210482635 dof: 392 pvalue: 0.04584767913524263
pspace_ks: scores: 196 D: 0.10042705348936298 pvalue: 0.03572416809095569
Test Problem bas-01
Surface Tally 11 - Z binned

tally_11_z_binned
\chi^2=22.73479 \text{ dof}=20 \text{ pval}=0.30200 N=196

-2.0 -1.5 -1.0 -0.5  0.0  0.5  1.0  1.5  2.0
0 25 50 75 100 125 150

Observed
Ideal
Test Problem bas-01
Surface Tally 11 - $P$ binned
Test Problem bas-01
Surface Tally 11 - Kolmogorov Smirnov
Test Problem bas-01
Surface Tally 11 - Kolmogorov Smirnov

The diagram shows a plot of the Kolmogorov-Smirnov test for the given tally. The values for the test are:
- Values: 196
- Maximum P Value: 0.10043
- P Value: 0.03572

The graph compares observed CDF, linear CDF, and Max D against the expected CDF.
### Basic Source in a Sphere

<table>
<thead>
<tr>
<th>IMP</th>
<th>N</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
<td>IMP:N=1 $ inside sphere</td>
</tr>
<tr>
<td>999</td>
<td>0</td>
<td>IMP:N=0 $ outside world</td>
</tr>
</tbody>
</table>

```
1  SO  1

print
MODE N
NPS 1E7
PRDMP 2J +1

c
SDEF POS=0 0 0 RAD=D1 ERG=D2 TME=D3

c
SI1 0 1 $ RAD distribution
SP1 -21 2

c
SI2  S 21 22 $ ERG distribution
SP2 1.01 2
SP21 -3 $ Watt spectrum
SI22 L 1.1 1.5 2.0 $ Discrete lines
SP22 2 1 3.05

c
SI3 H 10 100 $ Pulse
SP3 0 1

c
F1:N 1 $ ERG and TME tallies
E1 1E-2 63ilog 10
T1 5 63i 110
```
Test Problem bas-01 modified

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total scores</td>
<td>257</td>
</tr>
<tr>
<td>Usable scores</td>
<td>196</td>
</tr>
<tr>
<td>Unusable scores</td>
<td>60</td>
</tr>
<tr>
<td>Unusable matching zero scores</td>
<td>1</td>
</tr>
<tr>
<td>Unusable matching nonzero scores</td>
<td>0</td>
</tr>
<tr>
<td>Ref scores with errors &lt;= 0.05</td>
<td>196</td>
</tr>
<tr>
<td>New scores with errors &lt;= 0.05</td>
<td>197</td>
</tr>
<tr>
<td>Errors outside crit ratio</td>
<td>3.0 count: 0</td>
</tr>
<tr>
<td>Minimum non zero error over all ref data</td>
<td>0.0004</td>
</tr>
<tr>
<td>Minimum non zero error over all new data</td>
<td>0.0004</td>
</tr>
</tbody>
</table>
Test Problem bas-01 modified

Minimum non zero error over all new data: 0.0004
zspace_binned: scores: 196 chisq: 20.61121148274366 dof: 20 pvalue: 0.4203251973217062
z_summed_test: chisq: 298.8060630841057 dof: 196 pvalue: 3.0624173630059094e−06
Failure! Z Summed pvalue < pcrit 3.0624173630059094e−06 0.005
zspace ks: scores: 196 D: 0.13856774091732865 pvalue: 0.000958291677962661
Failure! KS normal pvalue < pcrit 0.000958291677962661 0.005
pspace_binned: scores: 196 chisq: 26.04081632653061 dof: 19 pvalue: 0.12905571066264981
pspace_summed: scores: 196 chisq: 521.3030690394801 dof: 392 pvalue: 1.2590408950300071e−05
Failure! P Summed pvalue < pcrit: 1.2590408950300071e−05 0.005
pspace ks: scores: 196 D: 0.1034116844159888 pvalue: 0.0280643749049152
Test Problem bas-01 modified
Surface Tally 11 - \( Z \) binned

tally\_11\_z\_binned
\[ \chi^2 = 20.61121 \quad \text{dof} = 20 \quad \text{pval} = 0.42033 \quad N = 196 \]
Test Problem bas-01 modified
Surface Tally 11 - $P$ binned

```
tally_11_p_binned
chi2=26.04082  dof=19  pval=0.12906  N=196
```

- Observed
- Ideal
Test Problem bas-01 modified
Surface Tally 11 - Z Kolmogorov Smirnov
Test Problem bas-01 modified
Surface Tally 11 - $P$ Kolmogorov Smirnov
### Test Problem bas-01 – report option

<table>
<thead>
<tr>
<th>Binname</th>
<th>Ref Val</th>
<th>Ref Err</th>
<th>New Val</th>
<th>New Err</th>
<th>Z-Value</th>
<th>P-Value</th>
<th>CI Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>E0</td>
<td>1.495e−04</td>
<td>0.02590</td>
<td>1.489e−04</td>
<td>0.02590</td>
<td>−0.10979</td>
<td>0.91258</td>
<td>ValidCIs</td>
</tr>
<tr>
<td>E1</td>
<td>6.400e−06</td>
<td>0.12500</td>
<td>6.700e−06</td>
<td>0.12220</td>
<td>Both invalid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E20</td>
<td>1.520e−05</td>
<td>0.08110</td>
<td>1.520e−05</td>
<td>0.08110</td>
<td>Exact Nonzero</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E41</td>
<td>4.050e−05</td>
<td>0.04970</td>
<td>4.120e−05</td>
<td>0.04930</td>
<td>0.24479</td>
<td>0.80662</td>
<td>ValidCIs</td>
</tr>
<tr>
<td>E42</td>
<td>3.970e−05</td>
<td>0.05020</td>
<td>4.240e−05</td>
<td>0.04860</td>
<td>Valid new, invalid ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E62</td>
<td>1.021e−04</td>
<td>0.03130</td>
<td>1.021e−04</td>
<td>0.03130</td>
<td>ValidCI Exact Values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E74</td>
<td>1.777e−04</td>
<td>0.02370</td>
<td>1.618e−04</td>
<td>0.02490</td>
<td>−2.72811</td>
<td>0.00637</td>
<td>ValidCIs LargeZ</td>
</tr>
</tbody>
</table>
Future work

- Is $P$ sensitive enough?
- Added statistical tests
- Distribute with MCNPTools
- Variance reduction testing
- Particle physics testing