ACTU PS5841 Data Science in Finance and Insurance - Autumn 2023 (Y. Wang)
Assignment - 1
Assigned 9/14/23, Due 9/22/23

A derivative contract pays the sum of squared "z-scores" of the performance of the 11 SP500 sectors over the next year. Assume performance is independent among sectors for now, the payoff of this derivative can be modeled by a chi-square distribution with 11 degrees of freedom, estimation issues aside.

It's not unreasonable for an uninformed trader to use the expected payoff to predict the payoff of this derivative. For some reason, you have trouble deriving the expected payoff and decided to do it numerically. You relied on the python `scipy` pacakge to produce a large sample from the $\chi^2(11)$ distribution, calculated the sample mean, and constructed a 95% confidence interval.

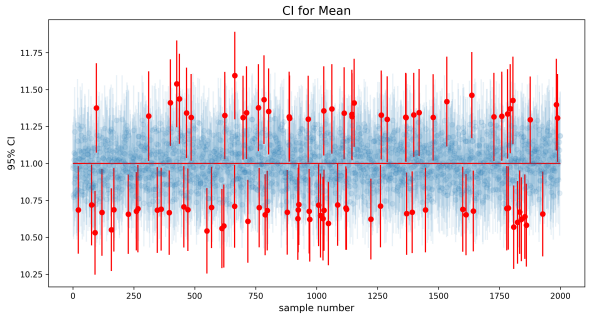
Your boss finally came to check your work and reminded you that the mean is exactly 11 in this case, and commented, "What does your confidence interval mean? Why don't you validate the concept for me numerically as well?"

Code in Python to calculate and validate the 95% confidence interval for the expected payoff.

- Do NOT use loops in your code.
- Draw 2000 samples, each of size 1000 from the $\chi^2(11)$ distribution.
- Construct the 95% confidence interval for the population mean for each sample.
- Calculate the proportion of these confidence intervals which do not contain the true average payoff (which is 11 according to your boss), and report these "outside" intervals by their indices.

```
94 intervals (or  4.7 %) does not contain the population mean
[  22   77   91   96  119  158  168  227  260  267  311  346  362  395
  400  426  437  455  466  472  485  550  569  611  619  623  663  664
  698  712  718  761  764  784  789  797  803  879  887  889  923  924
  927  965  969  972 1008 1016 1026 1029 1030 1048 1062 1086 1113 1120
 1122 1144 1145 1154 1222 1262 1265 1289 1365 1367 1369 1392 1398 1420
 1446 1478 1534 1599 1613 1636 1643 1727 1759 1780 1783 1786 1794 1805
 1808 1824 1833 1840 1854 1860 1876 1927 1983 1988]
```

- Present these confidence intervals graphically by giving those that do not contain the true average payoff a different color.



Please submit (*.ipynb and *.html) for your work

Sector Breakdown

SECTOR	INDEX WEIGHT
Information Technology	28.2%
Health Care	13.2%
Financials	12.5%
Consumer Discretionary	10.6%
Communication Services	8.8%
Industrials	8.4%
Consumer Staples	6.6%
Energy	4.4%
Materials	2.5%
Real Estate	2.4%
Utilities	2.4%

The weightings for each sector of the index are rounded to the nearest tenth of a percent; therefore, the aggregate weights for the index may not equal 100%.
Based on GICS® sectors
As of Aug 31, 2023