Computational Physics (Physics 760) Exercise #1

Dr. Evan Berkowitz `e.berkowitz@fz-juelich.de`
Dr. Stefan Krieg `s.krieg@fz-juelich.de`

Posted 18 Oct 2024
Due 25 Oct 2024 at 18:00

0 Course Setup

0.1 Make Buddies

If you haven't done so yet, take this opportunity to find a partner to do your homework with. We prefer teams of 2.

0.2 Version control with `git`

Once you have your team, please set up a `git` repository. Make sure you:

- Can use `git` on your own machine, or wherever you prefer to code.
- Understand and can use the following `git` commands:
  - help, init, status, diff, show, log
  - branch, checkout, merge
  - add, commit, push, pull

  There are many `git` tutorials online, in both written and video form.
- Add a file called STUDENTS with your names in it, one per line. Commit the file with a helpful message and push it.
- *Provide access to your tutor!*

It may be that you want to create one directory per assignment. It's not a bad idea. Just be aware that some of the assignments will be related to one another and you will want to reuse code (and maybe even some numerical results!).

0.3 Get access to JUSUF if you need it.

You can find details in the slides from 9 October, in the sciebo/Lecture notes/01.pdf.

1 Independent Trials + Trial Size [20 points]

Show business is part of scientific communication.
Think about how you present your results; you worked hard to create them.
Make sure your figures are clear and clearly explained!

Recall in lecture we had a method for stochastically computing π (or, equivalently, the area of a circle of unit radius). We did X eXperiments where we drew P Pairs of uniformly-distributed random numbers in the square of side length 2, $[-1, +1] \times [-1, +1]$.

On each pair we computed the indicator observable $[x^2 + y^2 \leq 1]$ which is 1 when the pair is inside the circle and 0 when it is outside. We averaged all P pairs in a single experiment labeled x

$$\pi_x = \left\langle 4[x^2 + y^2 \leq 1] \right\rangle_P = \frac{1}{P} \sum_p 4[x_p^2 + y_p^2 \leq 1] \tag{1}$$

and got X independent estimates for π, one π_x per $x \in X$.

We mentioned that we could take the X estimates (that is, repeat the estimation of π_x X different times using different random numbers each time) and average them to get a final answer, and take their standard deviation to quote an uncertainty.[1]

$$\text{final answer} = \frac{1}{X} \sum_x \pi_x \qquad\qquad \text{uncertainty} = \sqrt{\frac{1}{X-1} \sum_x (\pi_x - \text{final answer})^2}. \tag{2}$$

Obviously, as we increase either X or P we get a more reliable answer. In this exercise we will understand what exactly becomes more reliable as we increase these variables.

1.1 Just do one big experiment

Let $P = 10000$ and $X = 1$. Since $X = 1$ the uncertainty comes out to be 0 (or maybe ∞?). That poses a puzzle because of course you don't *actually* believe you got all the digits of π right (or wrong; certainly you learned *something*!). In this case sometimes people quote the standard deviation of the samples themselves. Let's try.

- Compute the mean and standard deviation of the π observable inside this single experiment.
- Plot a histogram of all the radii $r_p = \sqrt{x_p^2 + y_p^2}$ for each pair p.
- Plot a histogram of all the squared radii $r_p^2 = \sqrt{x_p^2 + y_p^2}$ for each pair p.
- Write a few sentences (which may include mathematics) explaining why the histograms have the features they do to the left of 1 and to the right of 1. In particular explain why the behavior differs so dramatically for values less than 1.
- Plot a histogram of the indicator variable $4[x^2 + y^2 \leq 1]$. Draw a vertical line at the mean of all the samples, and indicate the mean $\pm$ standard deviation with vertical lines. Indicate the true, known value of π for comparison.

1.2 Split into 100 experiments

Let's use the same number of random numbers in a different way. Let $P = 100$ and $X = 100$.

- Compute the estimate π_x for each experiment x. Compute the mean and its uncertainty using these estimates, as in (2).

[1] Maybe the denominator inside the uncertainty should be X; it makes very little difference in practice; don't worry about this for now.

- Plot a histogram of the estimates (there should be $X = 100$ numbers that go into your histogram, each one an average of $P = 100$ numbers). Indicate the true known value of π, and indicate the mean, and the mean $\pm$ the uncertainty.

1.3 A Zillion Little Experiments

What do we get if we try $P = 1$ and $X = 10000$? In other words, every patient is their own clinical trial. Are the results familiar? Explain.

1.4 Stop and think

Hopefully you found a standard deviation of about 1.6 for the first case (part 1.1) and an uncertainty of about 0.16 for the second case (part 1.2), even though you used the same number of random pairs, $XP = 10000$ in each case. (Of course, the exact values may differ a bit, because it's a randomized computation!)

- Were the estimates of the previous parts compatible with the known value of π?
- Write a few sentences explaining whether the standard deviation of a single experiment (as in 1.1) makes sense as an uncertainty.
- Write a few sentences explaining why just reorganizing the way we 'spent' our random numbers matters.

1.5 More Experiments vs. Longer Experiments

Let us investigate how the uncertainty behaves as we increase the number of experiments X and the length of each experiment P.

For all 16 possible combinations of $P \in (10, 100, 1000, 1000)$ and $X \in (10, 100, 1000, 10000)$:

- Get a final estimate of π and quote an uncertainty as in (2)
- Make a histogram of all the experimental results π_x. Draw the final estimate and the uncertainties around it.[2]
- If you can only afford a fixed number of random pairs across all experiments XP, how should you spend them? Use the 16 results to explain what you mean.

Now, let's make two more plots. Using all 16 uncertainties,

- For each P plot the uncertainty as a function of X; plot it as a log-log plot (with 4 lines or sets of points or however you like to draw it, one for each P; make sure the viewer can understand which is which!).
- For each X plot the uncertainty as a function of P; plot it as a log-log plot (with 4 lines or sets of points or whatever, one for each X; make sure the viewer can understand which is which!).
- Write a few sentences explaining what you see.

[2]Tip: some plotting utilities will 'helpfully' adjust the axis to zoom in on a narrow distribution, but we're exactly interested in comparing different distributions, so this helpful feature may make your point less clear.