

University of Kentucky, Physics 335
Laboratory #3, Rev. A, due Wednesday, 2023-09-13

This lab explores the distribution of the random variable X representing the number of fish caught in 100 minutes, assuming a 5% chance of catching a fish per minute, using Prof. Straley's simulation https://www.pa.uky.edu/~crawford/phy335_fa23/poisson.htm. Draw your plots using Python in Google Colab or another IDE following the example X01.ipynb from Laboratory #1.

1. Theoretical estimation of the Parent Distribution

a) What is the expected number λ of fishes to be caught in 100 min? Tabulate the Poisson distribution given λ for $x = 0, 1, 2, \dots, 10$.

b) Draw step plot of the Poisson distribution $P(x)$ out to $x = 10$, labeling your axes in the figure, which should span the entire width of your sheet. Plot the binomial distribution for $n = 10$, $np = \lambda$ (X02) on the same graph with the symbol $\times$. Why are the two distributions slightly different? [bonus: plot the binomial distribution for $n = 100$, which was used in the actual simulation.]

c) Calculate the $\mu = \lambda$ and $\sigma = \sqrt{\lambda}$ from the distribution. Why are they slightly different? Plot the mean as a vertical line on the distribution and indicate the $\mu \pm \sigma$ interval in each direction.

2. Experimental measurement of a Sample Distribution

a) Perform an experiment to estimate $P(x)$ by running the simulation $N = 25$ times, recording the numerical value of each result in a list. Calculate the mean $\bar{x}$ and standard deviation s of this sample. Make a new plot of each point $x_i \pm \delta x_i$ with error bars versus i on the abscissa. Draw a horizontal line at height $\bar{x}$ running through all points. How many error bars touch the line?

b) Add the frequency of occurrences of each value of x to a new line of the table in 1a). Normalize the distribution to 1 and plot $(h_x \pm \sqrt{h_x})/N$ with a solid circle and error bars in each bin on the graph of 1b). Draw a horizontal error bar representing the mean $\bar{x} \pm \delta \bar{x}$.

c) Calculate *chi-squared* $\chi^2 = \sum_x \left(\frac{h_x - NP(x)}{\delta h_x} \right)^2 = \sum_x \frac{(h_x - NP(x))^2}{h_x}$. Since terms with $h_x = 0$ are undefined, and bins with $h_x < 5$ have poor statistics, group x bins together to obtain super-bins of at least $\sum h_x \geq 5$ for each term in χ^2 . What is the average normalized deviation in each bin?

3. Calculation of the Combined Distribution

a) Tabulate your frequency distribution, mean, and standard deviation on the white board along with each other group. Calculate the combined sample distribution, its mean, and standard deviation $\bar{s}$ on a new line in the table of 1a).

b) Add the combined distribution $(h_x \pm \sqrt{h_x})/N$ from 3a) to the plot of 1b) using unfilled square makers with error bars. Plot the mean $\bar{x}_i$ from each group with ticks on the horizontal error bar of 2b). Add a horizontal error bar for the combined mean $\bar{x} \pm \delta \bar{x}$ above the previous one. Recalculate χ^2 . How did it change with higher statistics?

c) Qualitatively compare the plots and statistics of each of these distributions.