PHYS 275 – Experiment 2
Of Dice and Distributions

Experiment Summary

- Today we will study the distribution of dice rolling results
  - Two types of measurement, not to be confused: frequency with which we obtain a certain roll, and frequency with which we obtain a certain score
    - They answer to the following type of questions: how often do I get my die to show the face "2"? If I roll two dice, how often do I get a result of 10 or larger?
- We will study the probability distribution of dice rolls, and compare them with the expected theory
  - We can (and will!) calculate analytically the theoretical probability distribution
  - The comparison between theory and experimental results will tell us clues about our dice (are they fair?)

What should we learn today?

- Learn concepts of probability
  - What is probability?
  - How can I quantify random phenomena?
  - Power of probability: give indications on the outcome of a specific event
- Practice uncertainty analysis
  - Average of measurement, standard deviation, uncertainty on averages (see that repeating measurements you can make a better determination of a quantity)
  - Error propagation
  - Test of χ^2
- Learn how to present your data
  - Organize your spreadsheet in a way that it is easier for the TA to understand your analysis and results

Theory of Dice Rolling

- Let us start with 1 die
- If the die is fair, each side has equal probability of appearing, at each roll
  - Let us indicate with $P(n)$ the probability that the die will show the face with *n* dots
- We find that $P(n) = 1/6$ exploiting two trivial properties:
  - $P(1) = P(2) = \ldots = P(6)$: probability of falling on each side is the same, whichever the side
  - $\sum_{n=1}^{6} P(n) = 1$: the probability that the die falls on any side is 1; i.e., the die rolls, with probability 1, a number between 1 and 6
- The (discrete) function $P(n)$ is the <u>parent distribution</u>
  - It is the theoretical probability distribution for one die

Using the parent distribution

- The parent distribution is useful to define many properties of our system (one die)
  - Here are some of them, which you will encounter in this experiment
- Frequency distribution $f(n)$: number of times we see the n face
  - What is $\bar{f}(n)$, the <u>expected</u> number of times we see the n face if we throw the die N times?
    - $\bar{f}(n) = N * P(n)$

More on using $P(n)$

- Expected statistical fluctuation on $\bar{f}(n)$: $\sigma_{\bar{f}}(n)$
  - This has to depend on N: the more times you through the die, the closer your $f(n)$ should approach $\bar{f}(n)$
  - Formula: $\sigma_{\bar{f}}(n) = \sqrt{N * P(n) * (1 - P(n))}$
    - E.g.: if I throw my die 180 times, I expect to obtain $n = 6$ for $\bar{f}(n) \pm \sigma_{\bar{f}}(n) = 30 \pm 5$ times (try the math!)
    - Also note that since $P(n) = \frac{1}{6}$ (i.e., it is a constant), neither $\bar{f}(n)$ nor $\sigma_{\bar{f}}(n)$ actually depend on n: they are both constant
- Where is that formula coming out from?
  - Let us imagine that we threw a die N times, and want to study the probability of having it land on the n face for m times. What is the distribution of m? It is a Binomial distribution $B(m, N)$:
    - $B(m, N) = \binom{N}{m} P(n)^m (1 - P(n))^{N-m} = \frac{N! P(n)^m (1-P(n))^{N-m}}{m!(N-m)!}$

Yet more on using $P(n)$

- Expected average face value and standard deviation
  - This is the average score we expect to obtain if we throw our die. E.g., if I throw my die twice and obtain "4" and "1", my average face value is 2.5
- Let us throw our die N times
  - Since each side has equal probability of appearing, on average I will obtain "1" $N/6$ times, "2" $N/6$ times, and so on. Then, what is the expected average result?
    - $\langle n \rangle = \sum_{n=1}^{6} n * P(n) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = 3.5$

And yet more on using $P(n)$

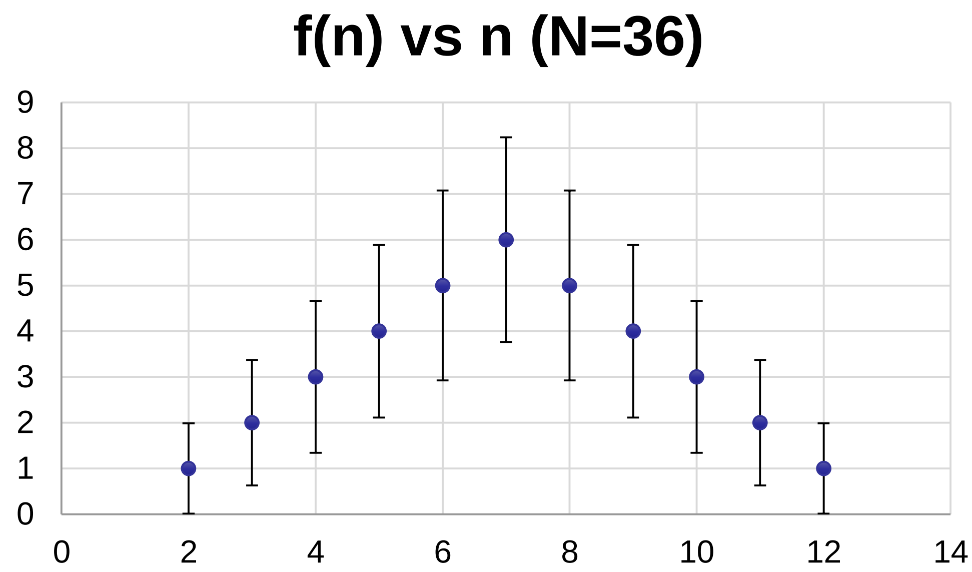
- What is the standard deviation?
  - Let us use its definition:
    - $\sigma^2 = \langle n - \langle n \rangle \rangle^2 = \sum_{n=1}^{6} (n - \langle n \rangle)^2 * P(n) = 2.917$
- Finally, if I throw a die once, what is the average score I expect to obtain?
  - $\langle n \rangle \pm \sigma = 3.5 \pm 1.7$
- What is the uncertainty on the average if I throw the die many times?
  - $\sigma_{\langle n \rangle} = {}^{\sigma}/_{\sqrt{N}} = {}^{1.7}/_{\sqrt{N}}$
- Please read the manual carefully
  - Do not confuse when you are asked to measure the frequency distribution and the average face value

Rolling two dice

- Now $P_{2dice}(n)$ gets complicate: n varies between 2 and 12, and its probability distribution is not flat anymore
  - It is more likely to obtain a score of 6 (1+5, 2+4, 3+3, 4+2, 5+1) than 2 (1+1 only)
- One can sit down and build a table $P_{2dice}(n)$:
  - $n = 2$: 1+1; one combination out of 36 achieves this result
  - $n = 3$: 1+2, 2+1; two combinations out of 36 work
  - $n = 4$: 1+3, 2+2, 3+1; three combinations out of 36 work
  - …

More on two dice

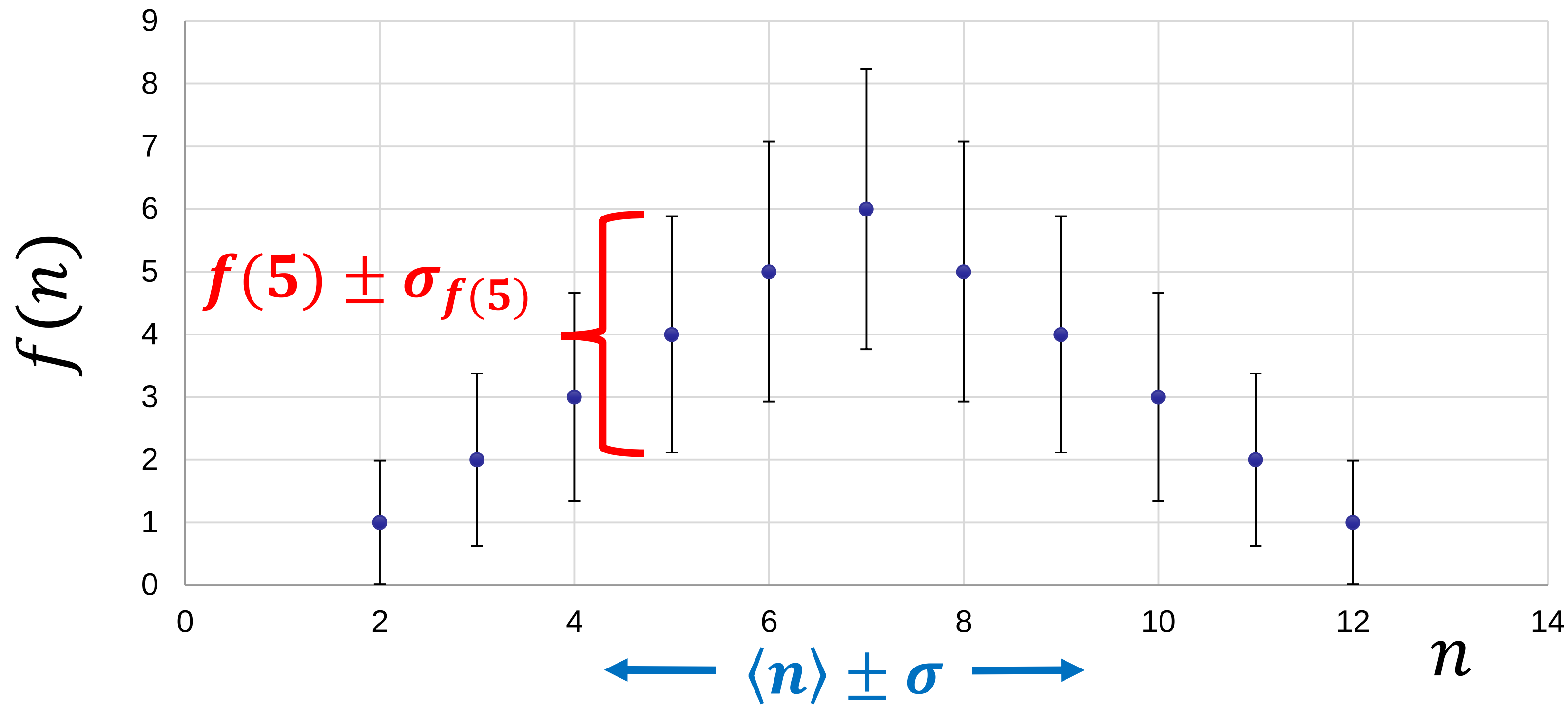
- It is easy to convince ourselves that we get:
  - $P_{2dice}(2) = P_{2dice}(12) = \frac{1}{36}$
  - $P_{2dice}(3) = P_{2dice}(11) = \frac{2}{36}$
  - $P_{2dice}(4) = P_{2dice}(10) = \frac{3}{36}$
  - $P_{2dice}(5) = P_{2dice}(9) = \frac{4}{36}$
  - $P_{2dice}(6) = P_{2dice}(8) = \frac{5}{36}$
  - $P_{2dice}(7) = \frac{6}{36}$



Last slide on two dice

- What is the average face value of two dice?
  - $\langle n \rangle_{2dice} = \sum_{n=2}^{12} n * P(n) = 7$
  - $\sigma_{2dice} = \sqrt{2}\sigma_{1die} = 2.415$
- What is the expected frequency distribution?
  - I.e., if I throw the two dice N times, how many times do I expect to see the score n?
  - $\bar{f}_{2dice}(n) = N * P_{2dice}(n)$
  - $\sigma_{\bar{f}_{2dice}}(n) = \sqrt{N * P_{2dice}(n) * (1 - P_{2dice}(n))}$
    - Note that now $P_{2dice}(n)$ does depend on n, hence $\bar{f}_{2dice}(n)$ and $\sigma_{\bar{f}_{2dice}}(n)$ also depend on n

$f(n) \pm \sigma_{f(n)}$ vs $\langle n \rangle \pm \sigma$



- Frequency
  - $f(n) = N * P(n)$
  - $\sigma_{f(n)} = \sqrt{N * P(n) * (1 - P(n))}$
- Score
  - $\langle n \rangle = \sum_n n * P(n)$
  - $\sigma = \sqrt{\sum_n (n - \langle n \rangle)^2 * P(n)}$

What will we be doing today?

- We reviewed what we expect to obtain when we throw a die or two dice
  - Let us do the experiment, and compare the theory with our experimental results
- How do we compare our results?
  - We perform a χ^2 test: $\chi^2 = \sum_n \left[\frac{f_{data}(n) - \bar{f}(n)}{\sigma_{\bar{f}}(n)}\right]^2$
    - Careful with n: from 1 to 6 in the 1-die case, from 2 to 12 in the 2-dice case
  - Please read discussion on degrees of freedom in the manual
- What if we obtain a result that has a very low probability of happening?
  - We are authorized to be suspicious. Today you will receive a die that may be loaded; find a way to decide if it is fair or not

Suggestions

- Plot B1
  - Make sure to put the theoretical distribution on the same graph as your measured distribution
  - The manual suggest to use “Data Analysis” to build the distribution of die rolls; I prefer to use the formula COUNTIF(<range>,<value>)
  - Use a Line/Scatter plot (please no Bar plots!)
- Questions B1-B3
  - Make sure it is clear what a χ^2 represents, and how to interpret $P(\chi^2, \nu)$
- Part D
  - Decide how to choose N, the number of times you need to throw a die to decide whether it is fair or not

More suggestions

- Questions B2, C3, C4, D1 (the χ^2 test)
  - Prepare a table similar to the following to calculate χ^2:

Experimental Die		Theory		χ^2	
Bin	Measured $f(n)$	$\bar{f}(n)$	$\sigma_{\bar{f}}(n)$	$(f(n)-\bar{f}(n))^2$	$\frac{(f(n)-\bar{f}(n))^2}{\sigma_{\bar{f}}(n)^2}$

- Running average in Excel:
  - AVERAGE(B1:B1): keep first cell of range fixed, and increase the other range delimiter

Today's experiments

- Probability distribution of one fair die
- Probability distribution of two fair dice
  - Roll and collect data of two dice together, then analyze the behavior of the first die, and the behavior of the two dice
- Probability distribution of a probably loaded die
  - Discuss with us before moving on to this part
- Important reminders
  - Submit your Excel spreadsheet on ELMS and turn in your check sheet before leaving the lab
  - Complete the final version of your report by 1pm next week
  - Finish the homework set in Expert-TA by 2pm next week
  - Save your data on the local disk frequently!

Critical Points

- Make sure you do roll your dice before leaving the laboratory
  - In particular, you need to roll a weird dice in part D: make sure to roll that dice before leaving the laboratory
    - If you have not reached part D by 5:30pm, skip questions B and C, jump to D and roll the dice, submit your draft report, then complete B and C in your final report
- Pay attention to what the manual asks you to do
  - Main source of confusion: sometimes you are counting how many times the die shows a certain face, and sometimes you are counting the score
- Make a clear distinction between averages, standard deviations, uncertainty on averages