

An Introduction to Statistics in Physics

As you see in the experiments, the arrival of an atom at a measurement counter is a random process. We would like to use the results of the experiments to determine the probability P that governs that random process. In the cases where all the atoms exit one port, then it is clear that the probability is 1 for that output state and zero for the other. However, if we measure 3 spin up atoms and 7 spin down atoms, then we must apply statistical analysis to help us solve the problem. Of course, those results would lead you to conclude that the probability of spin up is $P(+)=0.3$ and the probability of spin down is $P(-)=0.7$. However, if you performed the experiment a second time and counted 4 spin-up atoms and 6 spin-down atoms, then you would want to revise your estimates.

The questions we thus wish to address are: What is the best estimate of the probability, given the experimental data, and how confident are we of that estimate?

To answer these questions, let's first discuss what result we expect to obtain if we know the probability. Assume that a random process is governed by a probability P , and that each event is independent of all other events. Now assume that we have M of these events and we count the number of successes (e.g. spin-up atoms), which we call n . The probability that we count n spin-up atoms out of M total atoms is determined by the binomial probability distribution, and is given by

$$f_M(n) = \frac{M!}{(M-n)!n!} P^n (1-P)^{M-n}$$

This probability distribution is shown in Fig. A1 for the case $M=10$ and $P=0.5$. Thus, for

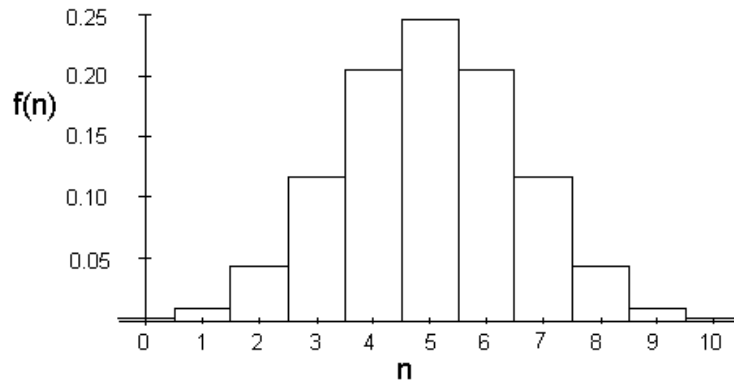


Fig A1. Binomial distribution for 10 events.

example, you expect to count 3 spin-up atoms about 12% of the time ($f_{10}(3) = 0.12$) and 5 spin-up atoms 25% of the time ($f_{10}(5) = 0.25$) in this case. The most obvious conclusion is that one single measurement of 10 atoms is not too reliable a predictor of the probability P than an atom is measured to have spin-up.

To reliably predict the probability we must perform repeated experiments and produce an experimental histogram of the data akin to the plot in Fig. A1. From the statistical properties of the histogram we can then estimate the probability and determine an error or uncertainty in that probability.

We generally characterize a probability distribution by 2 quantities: (1) the average or mean or expectation value, which is denoted by \bar{n} or $\langle n \rangle$, and (2) the standard deviation σ , which is the square root of the variance σ^2 . The mean tells you where the distribution is centered and the standard deviation tells you about the width of the distribution. The mean is obtained as a weighted average of the possible results:

$$\bar{n} = \sum_n n f(n)$$

where $f(n)$ is the probability of recording n counts. The variance is defined as

$$\sigma^2 = \sum_n (n - \bar{n})^2 f(n)$$

For the binomial distribution, the mean is

$$\bar{n} = MP$$

and the standard deviation is

$$\sigma = \sqrt{MP(1 - P)}$$

Experimental data is also commonly characterized by these two quantities. Consider an experiment where a variable x is measured N times to yield a data set x_i . The mean \bar{x} (or average value) of this data is

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

The standard deviation s of the data is

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N x_i^2 - \frac{N}{N-1} \bar{x}^2}$$

To connect this firmly to our experiments, assume that the variable x represents the number of times a certain result was obtained in M tries (e.g., M atoms leave the oven and we measure how many end up as spin-up). You would thus expect (and it is true) that the best experimental estimates of the parameters \bar{n} and σ of the theoretical distribution are the experimental parameters \bar{x} and s . Thus the experimental estimate of the probability of obtaining the desired result (e.g., the spin-up result) is

$$P = \frac{\bar{x}}{M}$$

What then is our uncertainty in this estimate? The first guess is to use the standard deviation of the data (divided by M to get a probability) since it is an estimate of the standard deviation of the theoretical probability distribution. However, this is not correct. The standard deviation of the data (and the theoretical

probability distribution) tells us how the data are distributed about the mean. The best estimate of the uncertainty of the mean, often called the standard deviation of the mean, is

$$\sigma_m = \frac{s}{\sqrt{N}}$$

which, as you might expect, tells us that we get a better estimate of the mean if we repeat the experiment more times.

A simple example may help to make this all more concrete. Consider an experiment where $10(M)$ coins are flipped and the number of heads (x) are counted, and the experiment is repeated 100 times (N). Figure A2 represents data from the experiment. The bars of the histogram tell us how many times a given number of heads occurred. The solid circles (connected by a solid line only as a guide to the eye) are the expected values given that the probability of a heads is $\frac{1}{2}$; this is just the binomial distribution shown in Fig. A1. The data have a mean of 5.42, with a standard deviation of 1.70, which you can see gives a measure of the width of the distribution of measurements but is much larger than what you might guess is the uncertainty of the mean value. (Note that if we do more experiments (increase N), the standard deviation s will not decrease, but we expect our uncertainty in the mean (i.e. the standard deviation of the mean) to decrease.) From this data we would estimate the probability P of a head and its uncertainty σ_p to be

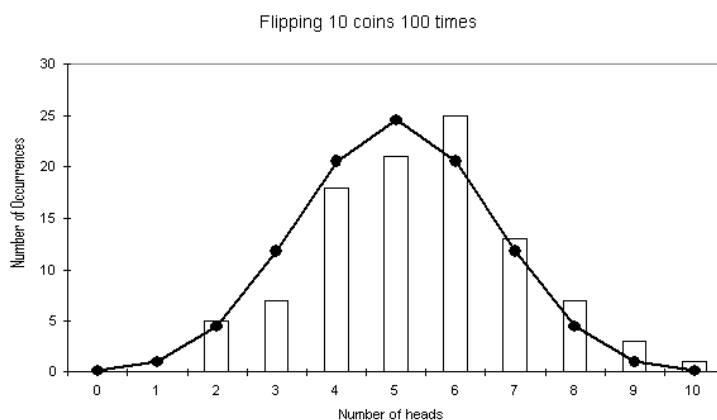


Figure A2: Experimental histogram of coin flipping

$$P = \frac{\bar{x}}{M} = \frac{5.42}{10} = 0.542$$

$$\sigma_p = \frac{\sigma_m}{M} = \frac{s}{M\sqrt{N}} = \frac{1.70}{10\sqrt{100}} = 0.017$$

Note that the uncertainty is about 3% of the value of the probability. This is a common result in statistics: if you are measure something N times, you can generally determine it wit a precision of $\frac{1}{\sqrt{N}}$. We already saw this in the standard deviation of the mean. In our counting experiments here, we are actually counting NM atoms and it shouldn't matter whether we measure them as N groups of M or M groups of N , or any

other combination; it's all the same data. This is evident if we recall that the standard deviation of the probability distribution scales as \sqrt{M} . Thus we expect the uncertainty in the probability to scale like:

$$\sigma_p = \frac{\sigma_m}{M} = \frac{s}{M\sqrt{N}} \propto \frac{\sqrt{M}}{M\sqrt{N}} = \frac{1}{\sqrt{MN}}$$

In the coin tossing example above $NM = 1000$ flips, so $\frac{1}{\sqrt{1000}} \approx 3\%$. In the 10 atoms case shown in #2 of the lab above, $NM = 100$ atoms, so $\frac{1}{\sqrt{100}} = 10\%$.

Note that the experimental estimate of the probability in the coin tossing example above differs from what we know the real value to be by about 2.5 times the standard deviation. This is only expected to happen 1.5% of the time, but it can happen. We expect our results to be within one standard deviation 68% of the time and within 2 standard deviations 95% of the time.

Activity: Determining if a Die is Fair

Each group will be given a die. (Make sure that before you leave class today, you mark your die in an identifiable way with a piece of masking tape, in case you need to use it again.) Design and execute an experiment to check if your die is fair.