Probability and Statistics

Due to the **probabilistic nature** of quantum mechanics, it is useful to know several basic concepts of probability and statistics.

Consider an experiment (such as rolling of a die) which has 'n' possible outcomes, each with a probability P_j.

- *The probability of outcome 'j' occurring is given by* P_j

With each experiment (or roll of the die) it is certain a result will occur. In other words there is a 100% probability of an event occurring. Thus, the sum of all probabilities is unity since one of the '*n*' outcomes must occur.

$$1 = \sum_{j=1}^{n} P_j$$

Consider a **4-sided die**. The probabilities are given in the table:

outcome	**probability**
1	0.25
2	0.25
3	0.25
4	0.25

total: 1.00

Now let's associate a value 'x_j' with each outcome.

For example, say with our 4-sided die we have numbers 1, 3, 7 and 10 on the sides, instead of 1, 2, 3 and 4.

outcome, j	**result, x_j**	**Probability, P_j**
1	1	0.25
2	3	0.25
3	7	0.25
4	10	0.25

The average or mean of the result 'x' is given the symbol $\langle x \rangle$ and is given by:

$$\langle x \rangle = \sum_{j=1}^{n} x_j P_j$$

In the specific example of our 4-side die above we have:

$$\langle x \rangle = (1)(0.25) + (3)(0.25) + (7)(0.25) + (10)(0.25)$$

$$\langle x \rangle = 5.25$$

The average result we will obtain is 5.25.

The **variance**, σ^2, indicates the extent to which the set of outcomes or observations cluster around the mean value.

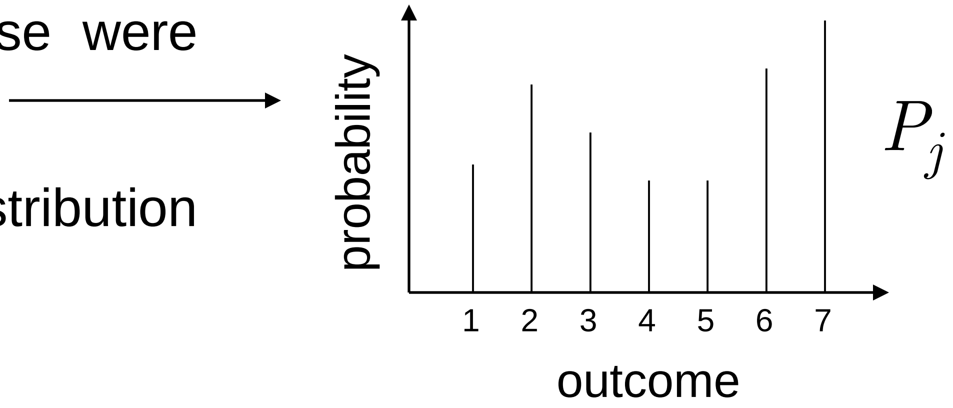
The variance is equal to the average of the squared deviations from the mean:

$$\sigma^2 = \sum_{j=1}^{n} P_j \left(x_j - \langle x \rangle \right)^2$$

The **standard deviation**, σ, is the square root of the variance, a measure of the width of its “spread”:

Discrete and Continuous Probabilities

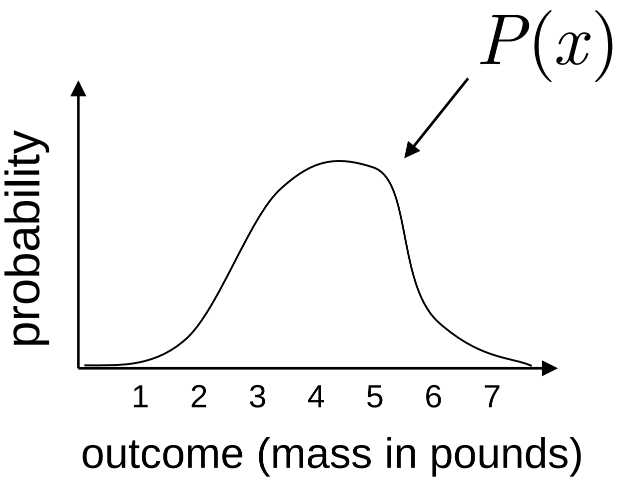
With the case of the dice, these were discrete probability distributions:



In this case the probability distribution function is discrete.

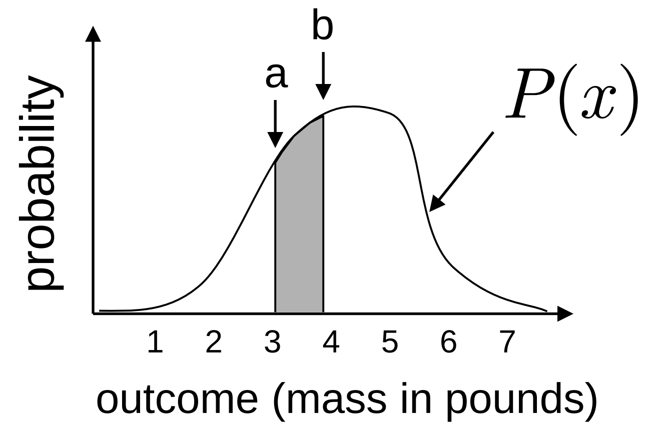
In this course we will more often have to deal with continuous probability distributions.

For example the mass of water melons in a large batch will be given by a continuous probability since the mass of the melons can be any value greater than zero.



Such a probability distribution can be described by a continuous function $P(x)$

In our example, the probability of finding a water melon between mass ‘a’ and mass ‘b’ is given by the integral:



$$P_{a-b} = \int_a^b P(x)dx$$

As with the discrete probabilities, the sum of all probabilities must sum to unity or 100%. Thus, our continuous distribution function must satisfy the following equation:

$$1 = \int_{\text{all possible values}} P(x)dx$$

In the case the function describing the mass of water melons we have:

$$1 = \int_0^\infty P(x)dx$$

In analogy to the discrete distributions, the average outcome given by a continuous probability distribution is:

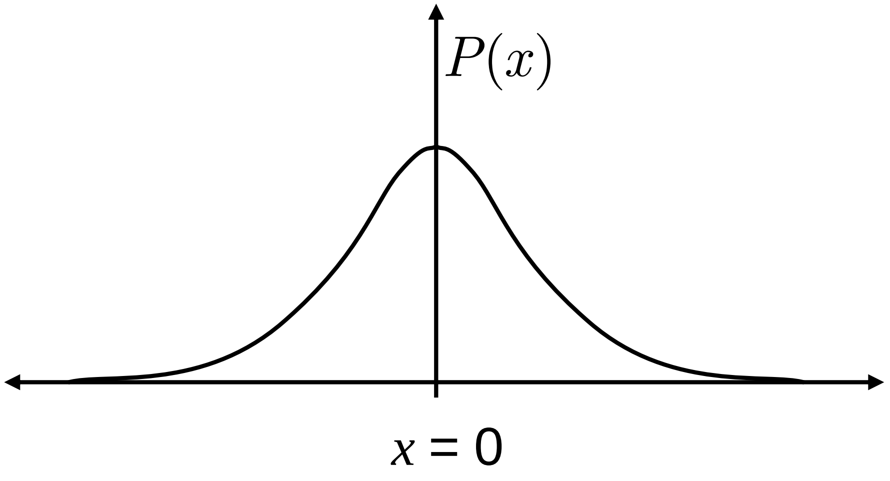
$$\langle x \rangle = \int x \cdot P(x)dx$$

Example

A very important probability distribution is a ***Gaussian*** distribution of random errors. σ^2 is the variance. σ is the standard deviation.

$$P(x) = \frac{1}{\sigma}\sqrt{\frac{1}{2\pi}}e^{-x^2/2\sigma^2}$$

$$-\infty < x < +\infty$$



Show that the probability distribution function 'sums' to unity:

$$\int_{-\infty}^{+\infty} P(x)\mathrm{d}x = \int_{-\infty}^{+\infty} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}\mathrm{d}x \qquad \text{with } \alpha = 1/(2\sigma^2)$$

$$= \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha x^2}\mathrm{d}x$$

$$= \sqrt{\frac{\alpha}{\pi}} \cdot \left(\frac{\pi}{\alpha}\right)^{1/2}$$

$$= 1$$

From a table of integrals this integral is:

$$\left(\frac{\pi}{\alpha}\right)^{1/2} = \int_{-\infty}^{+\infty} e^{-\alpha x^2}\mathrm{d}x$$

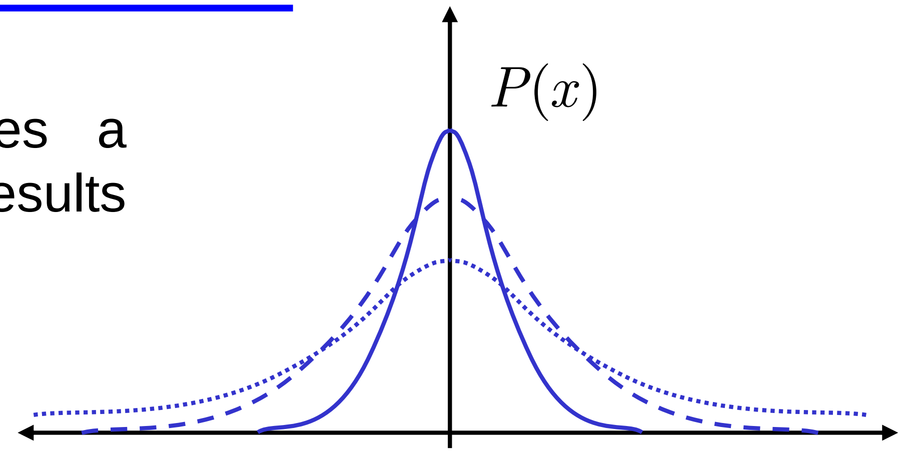
What integral would we have to evaluate to determine the average outcome and what should the average be?

answer: $$\langle x \rangle = \int_{-\infty}^{+\infty} x \cdot \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}\mathrm{d}x$$

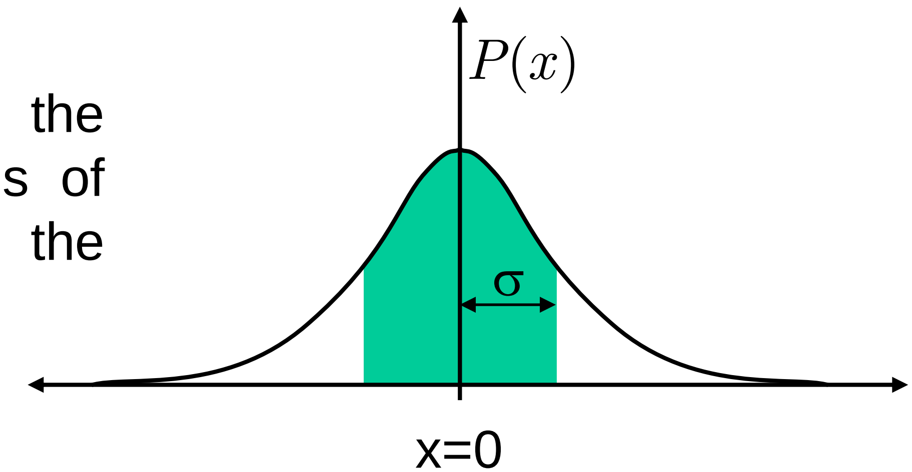
The average x value should be zero based on the symmetry of the probability around x = zero

Standard Deviation

The standard deviation, σ, gives a measure of how spread out the results are from the mean.



For a continuous distribution the standard deviation defined in terms of the variance, σ^2, in analogy to the discrete definition.



$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$$

$$= \int x^2 \cdot P(x)dx - \left(\int x \cdot P(x)dx \right)^2$$

Important Application to Quantum Chemistry

Suppose that the wave function of a system is ψ_n, an eigenfunction of operator $\underline{A}$ with eigenvalue a_n.

Then $\underline{A}\,\psi_n = a_n\,\psi_n$ and

$$<a> = \int \psi_n^* \underline{A}\,\psi_n d\tau = \int \psi_n^* a_n\,\psi_n d\tau$$

$$= a_n \int \psi_n^* \psi_n d\tau = a_n$$

Furthermore:

$$<a^2> = \int \psi_n^* \underline{A}^2\,\psi_n d\tau = \int \psi_n^* a_n \underline{A}\,\psi_n d\tau$$

$$= \int \psi_n^* a_n^2\,\psi_n d\tau = a_n^2 \int \psi_n^* \psi_n d\tau = a_n^2$$

The standard deviation in a is

$$\sigma_a^2 = <a^2> - <a>^2 = a_n^2 - a_n^2 = 0$$

»» **The only measured value is eigenvalue a_n**