Probability

If one gets outcome \( i \) \( N_i \) times in \( N \) tries, then the probability \( P(i) \) is the limit
\[
P(i) = \lim_{N \to \infty} \frac{N_i}{N}.
\]

Note \( 0 \leq P(i) \leq 1 \).

\[
P(i \text{ or } j) = P(i) + P(j) \quad \text{for } i \neq j.
\]

\[
\sum_{\text{all } i} P(i) = \sum_{\text{all } i} \frac{N_i}{N} = \frac{1}{N} \sum_{\text{all } i} N_i = \frac{N}{N} = 1.
\]

These are the rules for distinct discrete outcomes.
Now suppose that $x$ is a continuous real variable.

Suppose the possible values of $x$ are all real numbers $a \leq x \leq b$.

We can split the interval $[a, b]$ into $H$ bins of size

$$dx = \frac{b-a}{H}$$

where $H$ is a huge number.

We use the same definition where now the index $i$ labels the bins.

Bin 1 is the interval $[a, a+d x]$,

Bin 2 $\ldots$ Bin $H$ $\ldots$ $[a+H d x, b]$. 

We say

$$p(x) = \lim_{N \to \infty} \frac{N_i}{N}$$

but now these $p(i)$'s are proportional to the size $d x$ of the bins.
So to avoid reference to the width of the bins, we set for \( x \) in bin \( i \):

\[
P(x) \, dx = \left( \lim_{i \to x} \frac{N_i}{N} \right) = P(i).
\]

Now the probability distribution

\[
P(x) = \frac{P(i)}{dx}
\]

is no longer bounded by unity.

For instance, if all the \( x_i \)'s were to fall in the first bin, then \( P(1) \) would be 1 and \( P(x) \) would be

\[
P(x) = \frac{P(1)}{dx} = \frac{1}{dx} \text{ for } a \leq x \leq a+dx.
\]

and

\[
P(x) = \frac{0}{dx} = 0 \text{ for } x > a+dx,
\]

So \( P(x) \) would be big for \( a \leq x \leq a+dx \) if \( dx \) is small.
\[ \int_a^b P(x) \, dx = \sum_i P(i) = \frac{\sum_i N_i}{N} = \frac{\sum_i N_i}{N} = \frac{N}{N} = 1 \]

So the probability distribution \( P(x) \) is normalized to unity over the interval of possible outcomes.

The uniform distribution on \([a, b]\) is

\[ P(x) = \frac{1}{b-a} \quad \text{for} \quad a \leq x \leq b \]

\[ = 0 \quad \text{for} \quad x < a \text{ or } x > b. \]

The Gaussian distribution is

\[ P(x) = A e^{-\frac{(x-x_0)^2}{2\sigma^2}} \]

where \( A \) and \( \sigma \) are positive numbers

and \( x_0 \) is a possible value of the real variable \( x \).
The integral now runs from 

\(-\infty \to 0 \to \infty\).

\[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dxdy e^{-x^2-y^2} = \int_{-\infty}^{\infty} e^{-r^2} r d\theta = \pi \int_{0}^{\infty} e^{-u} du\]

\[= \pi \left[-e^{-u} \right]_{0}^{\infty} = \pi \left(0 - (-1) \right) = \pi\]

So, \[\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}\]
So to normalize the probability distribution

\[ P(x) = A e^{-\frac{(x-x_0)^2}{2 \sigma^2}} \]

we compute

\[ 1 = A \int_{-\infty}^{\infty} e^{-\frac{(x-x_0)^2}{2 \sigma^2}} \, dx \]

by setting \( y = \frac{x-x_0}{\sigma \sqrt{2}} \) so that

\[ dy = \frac{dx}{\sigma \sqrt{2}} \], \quad \text{Then} \quad dx = \sigma \sqrt{2} \, dy \quad \text{and} \quad \]

\[ 1 = A \sigma \sqrt{2} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \, dy = A \sigma \sqrt{2} \sqrt{\pi} \]

whence

\[ A = \frac{1}{\sigma \sqrt{2} \pi} \]

So

\[ P(x) = \frac{1}{\sigma \sqrt{2 \pi}} e^{-\frac{(x-x_0)^2}{2 \sigma^2}} \]

is the normalized probability distribution.
The mean is
\[ \langle x \rangle = \sum_i x_i P(i) \quad \text{discrete} \]
\[ = \int dx \cdot P(x) \quad \text{continuous} \]
The mean is the average. It is also called "the expected value" or "the expectation value," (sic).

The mean need not be the most probable value. For instance,

\[ -5 \quad \quad 5 \]

The most probable value would be \( x = 5 \), but the mean \( \langle x \rangle \) would be tiny, near \( x = 0 \), and the probability distribution \( P(x) \) would have the small value \( P(\langle x \rangle) \) at the mean \( \langle x \rangle \).
The mean of $f(x)$ is

$$\langle f \rangle = \langle f(x) \rangle = \sum_i f(x_i) \theta(i) \quad \text{discrete}$$

$$= \int dx \ f(x) \ P(x) \quad \text{continuous}.$$ 

If $f(x) = c$ is a constant, then

$$\langle f(x) \rangle = \langle c \rangle = \int dx \ c \ P(x) = c.$$ 

So

$$\langle \langle f \rangle \rangle = \langle f \rangle.$$ 

The rms deviation is

$$\text{rms} d = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$$

$$= \left( \int dx \ (x - \langle x \rangle)^2 \ P(x) \right)^{1/2}$$

where $\langle x \rangle = \int dx \ x \ P(x)$ is the mean.
If the rms deviation \((\text{rmsd})\) is zero, then the probability distribution \(P(x)\) must be concentrated at the mean \(<x>\).

\(P(x)\):

\(<x>\)

Some use Dirac's notation \(P(x) = \delta(x - <x>)\) for this case.

The square of the \(\text{rmsd}\) is the variance.

Note that

\[ (\text{rmsd})^2 = \int dx \ (x - <x>)^2 P(x) \]

\[ \text{variance} = \int dx \ (x^2 - 2x<x> + <x>^2) P(x) \]

\[ \text{variance} = <x^2> - 2<x>^2 + <x>^2 = <x^2> - <x>^2. \]
The uniform distribution $P(x)$ on the interval $[0, a]$ is 

$$P(x) = \frac{1}{a} \quad \text{for} \quad 0 \leq x \leq a$$

and $P(x) = 0$ otherwise. What is the variance or $\text{var}^2$? Since $P(x)$ is uniform, the mean must be $a/2$

$$\langle x \rangle = \int_0^a dx \cdot x \cdot P(x) = \int_0^a dx \cdot \frac{x}{a}$$

$$= \frac{1}{a} \int_0^a dx \cdot x = \frac{1}{a} \left[ \frac{x^2}{2} \right]_0^a$$

$$= \frac{a^2}{2a} = \frac{a}{2} \quad \text{as expected.}$$

So $\langle x \rangle = a/2$, and the variance is 

$$\text{var}^2 = \langle x^2 \rangle - \langle x \rangle^2.$$ We need $\langle x^2 \rangle$:

$$\langle x^2 \rangle = \int_0^a dx \cdot x^2 \cdot P(x) = \frac{1}{a} \int_0^a dx \cdot x^2 = \frac{1}{a} \left[ \frac{x^3}{3} \right]_0^a = \frac{a^2}{3}.$$
So the square of the variance is

\[ \text{var} = \langle x^2 \rangle - \langle x \rangle^2 = \frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12}. \]

What is the variance for the Gaussian distribution

\[ P(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}. \]

Clearly, the mean \( \langle x \rangle = x_0 \)

\[ \langle x \rangle = \int_{-\infty}^{\infty} dx \ x \ P(x) = \int_{-\infty}^{\infty} dx \ x \ e^{-\frac{(x-x_0)^2}{2\sigma^2}} \]

so we can prove by adding and subtracting

\[ x_0 \]

\[ \langle x \rangle = x_0 + \int_{-\infty}^{\infty} dx \ x \ e^{-\frac{(x-x_0)^2}{2\sigma^2}} - x_0 \]

\[ = x_0 + \int_{-\infty}^{\infty} dx \ \frac{x-x_0}{\sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\sigma^2}} \]

since \( \int_{-\infty}^{\infty} P(x) = 1 \).
Let \( u(x) = \frac{(x-x_0)^2}{2\sigma^2} \). Then we have

\[
\frac{du}{dx} = \frac{x-x_0}{\sigma^2}
\]

and so

\[
\langle x \rangle = x_0 + \int_{-\infty}^{\infty} dx \frac{1}{\sqrt{2\pi}} \frac{du(x)}{dx} e^{-u(x)}
\]

\[
= x_0 + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx u(x) e^{-u(x)}
\]

\[
= x_0 + \frac{\sigma}{\sqrt{2\pi}} \left[ -e^{-u(x)} \right]_{-\infty}^{\infty} = x_0.
\]

So \( \langle x \rangle = x_0 \) as expected.

We could compute \( \langle x^2 \rangle \),

but it is easier to compute the variance directly,

Since \( \mu_0 \) is the mean, the variance is by definition
\[
\text{variance}(x) = \int_{-\infty}^{\infty} (x - \mu_0)^2 \cdot P(x) \, dx
\]
\[
= \int_{-\infty}^{\infty} \frac{(x - \mu_0)^2}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{(x - \mu_0)^2}{2\sigma^2}} \, dx
\]

We now change variables as before.

Setting \( y = \frac{x - \mu_0}{\sigma \sqrt{2}} \) so that
\[
\int_{-\infty}^{\infty} \frac{-(x - \mu_0)^2}{2\sigma^2} \cdot e^{-\frac{(x - \mu_0)^2}{2\sigma^2}} \, dx = \int_{-\infty}^{\infty} -y^2 \cdot e^{-y^2} \, dy
\]

\[
(x - \mu_0)^2 = 2\sigma^2 y^2
\]

\[
dy = \frac{dx}{\sigma \sqrt{2}} \quad \text{and} \quad dx = \sigma \sqrt{2} \, dy
\]

Putting these together, we find
\text{variance } \langle x \rangle = \int_{-\infty}^{\infty} dx \, \frac{(x - x_0)^2}{2 \sigma^2} e^{-\frac{(x - x_0)^2}{2 \sigma^2}}

= \frac{1}{\sqrt{2 \pi} \sigma} \int_{-\infty}^{\infty} dy \, \sqrt{b} \, e^{-\frac{y^2}{2}} \left( \frac{2 \sigma^2}{b} \right)^{\frac{y^2}{2}}

= \frac{2 \sigma^2}{\sqrt{\pi} b} \int_{-\infty}^{\infty} dy \, y^2 e^{-y^2}

\text{Let } I(b) = \int_{-\infty}^{\infty} dy \, e^{-by^2}

\text{With } z = \sqrt{b} y, \quad dz = \sqrt{b} \, dy, \quad dy = dz / \sqrt{b}

\text{and } I(b) = \frac{1}{\sqrt{b}} \int_{-\infty}^{\infty} dz \, e^{-z^2} = \frac{\sqrt{\pi}}{\sqrt{b}} = \sqrt{\frac{\pi}{b}}

\text{since we already did the integral}

\int_{-\infty}^{\infty} dx \, e^{-x^2} = \sqrt{\pi}.
\[
\sqrt{\frac{\pi}{b}} = I(b) = \int_{-\infty}^{\infty} dy \ e^{-by^2}
\]

Differentiate both sides with respect to \(b\):

\[
\frac{d}{db} \sqrt{\frac{\pi}{b}} = \sqrt{\frac{\pi}{b}} \left( -\frac{1}{2} \right) b^{-3/2} = -\frac{1}{2} \sqrt{\frac{\pi}{b^3}}
\]

\[
= \frac{d}{db} \int_{-\infty}^{\infty} dy \ e^{-by^2} = \int_{-\infty}^{\infty} dy \ \frac{d}{db} e^{-by^2}
\]

\[
= -\int_{-\infty}^{\infty} dy \ y^2 e^{-by^2}
\]

So by this method, we get

\[
\int_{-\infty}^{\infty} dy \ y^2 e^{-by^2} = \frac{1}{2} \sqrt{\frac{\pi}{b^3}}
\]

On setting \(b = 1\),

\[
\int_{-\infty}^{\infty} dy \ y^2 e^{-y^2} = \frac{\sqrt{\pi}}{2}.
\]
The variance of the gaussian

\[ P(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\sigma^2}} \]

is

\[ \text{variance}(x) = \frac{2\sigma^2}{\sqrt{n}} \frac{\sqrt{n}}{2} = \sigma^2. \]

Since the variance is the \text{rmsd} squared, we also have

\[ \text{rmsd} = \sigma. \]

For the gaussian \( P(x) \),

probabilities for exclusive events add. The probability that \( x \) is either less than 1 or greater than 2 is

\[ P(x<1 \text{ or } x>2) = P(x<1) + P(x>2) \]

\[ = \int_{-\infty}^{1} dx \ P(x) + \int_{2}^{\infty} dx \ P(x). \]
But the probability that \( x > 1 \) or \( y > 2 \) need not be the sum \( P(x > 1, y > 2) = P(x > 1) + P(y > 2) \), because an event can have \( x > 1 \) and \( y > 2 \).

If two variables \( x \) and \( y \) are independent, then their joint probability distribution is the product \( P(x, y) = P(x) P(y) \) and it is properly normalized

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) \, dx \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x) P(y) \, dx \, dy = \int_{-\infty}^{\infty} P(x) \, dx \int_{-\infty}^{\infty} P(y) \, dy = 1 \cdot 1 = 1.
\]
Review: Discrete events

\[ P(i) = \lim_{N \to \infty} \frac{N_i}{N} \]

Continuous variable \( x \)

\[ P(x) \, dx = \lim_{N \to \infty} \frac{N_i}{N}, \quad N_i \text{ is number in } [x - \frac{dx}{2}, x + \frac{dx}{2}] \]

\[ 1 = \int P(x) \, dx = \sum_i P(i) \]

Gaussian distribution

\[ P(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x - x_0)^2}{2\sigma^2}} \]

has mean

\[ \langle x \rangle = \int dx \, x \, P(x) = x_0 \]

and variance

\[ \langle (x - x_0)^2 \rangle = \langle x^2 \rangle - x_0^2 \]

\[ = \int dx \, (x^2 - x_0^2) \, P(x) = \sigma^2 \]
Ex.

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

\[ P(y) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}} \]

\[ P(x,y) = P(x)P(y) = \frac{1}{2\pi \sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \]

Let \( r^2 = x^2 + y^2 \) then

\[ P(x,y) = \frac{1}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2}} \]

The area of the ring is \( 2\pi r \, dr \), so

\[ P(r) \, dr = \frac{1}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2}} \, 2\pi r \, dr \]

\[ = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \, dr \]

So probability \( x,y \) in ring at \( r \) of width \( dr \).
\[ P(r) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \]

\[ 6 = 5 + 1 = 1 + 5 = 4 + 2 = 2 + 4 = 3 + 3 = 5 \]

\[ P(6) = \frac{5}{6} = \frac{5}{36} \]

\[ P(12) = \frac{1}{36} \]

\[ P(n > R_0) = \int_{R_0}^{\infty} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} dr \]

Let

\[ u = \frac{r^2}{2\sigma^2} \]

\[ du = \frac{r}{\sigma^2} \, dr \]

\[ P(n > R_0) = \int_{R_0}^{\infty} du \, e^{-u} = \left[ -e^{-u} \right]_{R_0}^{\infty} \]

\[ = e^{-u(R_0)} - e^{-\frac{R_0^2}{2\sigma^2}} \]

\[ = e^{-u(R_0)} - e^{-\frac{R_0^2}{2\sigma^2}} \]
\[ P(v) = P(v_x) P(v_y) P(v_z) \]

\[ = \frac{-v_x^2}{2\sigma_x^2} \cdot \frac{-v_y^2}{2\sigma_y^2} \cdot \frac{-v_z^2}{2\sigma_z^2} \cdot e^{-\frac{v^2}{2\sigma^2}} \]

\[ = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-\frac{u^2}{2\sigma^2}} \]

where \( u^2 = v_x^2 + v_y^2 + v_z^2 \).

The volume of the shell is \( 4\pi u^2 \, du \), so

\[ \int P(u) \, du = \frac{4\pi u^2}{(2\pi\sigma^2)^{3/2}} e^{-\frac{u^2}{2\sigma^2}} \, du \]

\[ = \frac{2}{\sigma^3} \frac{u^2}{\sqrt{2\pi \sigma^2}} e^{-\frac{u^2}{2\sigma^2}} \, du \]

\[ = \left( \frac{1}{\sqrt{2\pi \sigma^2}} \right)^3 e^{-\frac{u^2}{2\sigma^2}} 4\pi u^2 \, du. \]
Ideal - Gas Law

"Ideal" means that the gas molecules do not interact with each other, except to scatter elastically.

\[ pV = NkT \] is the ideal gas law

where \( N \) is the number of molecules of gas in the volume \( V \) at pressure \( p \) and temperature \( T \) in K. Here

\[ k = 1.38 \times 10^{-23} \text{ J/K} \] is the Boltzmann constant

\[ hT_r \approx 4.1 \text{ pJ nm} \]

where \( T_r \approx 295 \) is room temperature.

Now \( [hT] = \text{force \times distance} = \text{N m} = \text{energy} \),

Also, \( pV = pA \) and \( [p] = \frac{\text{force}}{\text{area}} \), so \( [pA] = \text{force}, \text{ so } [pV] = \text{energy} \).
The time between impacts on the right wall is \( \Delta t = \frac{2L}{v_x} \). The momentum transferred to the wall by each impact is \( \Delta p_x = 2m v_x = 2p_x \).

So the average force due to one molecule of mass \( m \) is

\[
f = \frac{2m v_x}{\Delta t} = \frac{2m v_x}{(2L/v_x)} = \frac{mv_x^2}{L}.
\]
If \( A \) is the area of the right wall, then the pressure on it is
\[
p = \frac{f}{A} = \frac{m v_x^2}{L A},
\]

If the box is a cube, then \( L A = V \), and we have
\[
p = \frac{m v_x^2}{V} \quad \text{for one molecule.}
\]

For a cube with \( N \) molecules, all of mass \( m \), all moving at speed \( v_x \),
\[
p V = N m v_x^2.
\]

Since the molecules have different speeds, we take the average,
\[
\langle p V \rangle = p V = N m \langle v_x^2 \rangle.
\]
But the ideal-gas law is

\[ pV = NkT. \]

So

\[ pV = NkT = N\text{m} \langle v_x^2 \rangle. \]

So

\[ m \langle v_x^2 \rangle = kT. \]

Now

\[ v^2 = v_x^2 + v_y^2 + v_z^2. \]

and so

\[ \langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle. \]

but surely

\[ \langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle, \]

So

\[ \langle v^2 \rangle = 3 \langle v_x^2 \rangle \quad \text{and so} \]

\[ m \langle v^2 \rangle = 3kT. \]
Thus the mean kinetic energy of a molecule of mass \( m \) in an ideal gas at pressure \( p \) and temperature \( T \) is

\[
\frac{m}{2} \langle u^2 \rangle = \frac{3}{2} kT.
\]

A mole of air occupies 24 L or 0.024 m\(^3\) at atmospheric pressure \( p_0 \) and room temperature \( T_0 \).

At sea level, \( p_0 \) can lift a column of water about 10 m.

The force needed to support that column is

\[
f = 10 \text{ m} \cdot p_{\text{H}_2\text{O}} A g
\]

where \( A \) is the area of the column.
Now \( p_{\text{H}_2\text{O}} = \frac{1.9 \text{m}^3}{\text{cm}^3} = \frac{10^{-3} \text{kg}}{(10^{-2} \text{m})^3} = \frac{10^3 \text{kg}}{\text{m}^3} \).

So,

\[ \rho = \frac{f}{A} = 10 \text{m} \times \frac{10^3 \text{kg}}{\text{m}^3} \]

\[ = 10^4 \frac{\text{kg}}{\text{m}^2} \times 9.8 \text{N} = 10^5 \frac{\text{N}}{\text{m}^2} \]

\[ = 10^5 \text{ Pa} \]

where \( \text{Pa} = \frac{\text{N}}{\text{m}^2} \) is one Pascal.

So, for one mole of air

\[ pV = NkT \]

and \( kT = 4.1 \text{ pN m} \times 10^{-12} \text{ N m} \]

\[ = 4.1 \times 10^{-21} \text{ J} \]

and \( NkT \approx 6.0 \times 10^{-23} \times 4.1 \times 10^{-21} \text{ J} = 2460 \text{ J} \).
while one mole of air occupies $24\,L = 0.024\,m^3$ (at STP)

so

\[ pV = 10^5 \frac{N}{m^2} \times 0.024\,m^3 = 2400\,J \]

so this works to the precision of the values we are using.

The molar mass of $N$ is $14\,g$.

so one mole of $N_2$ is $28\,g$. So

The mass of a single molecule of $N_2$ is

\[ \frac{14\,g}{N_a} = \frac{28 \times 10^{-3}\,kg}{6 \times 10^{23}} = 4.7 \times 10^{-26}\,kg \]

\[ \frac{1}{2} m \langle v^2 \rangle \geq \frac{3}{2} kT \leq \frac{3}{2} \times 4.1 \times 10^{-21} \,J \]

\[ \langle v^2 \rangle = 12.3 \times 10^{-21} \,J \geq \frac{12.3 \times 10^{-21}}{4.7 \times 10^{-26}} \,kg \]

\[ v = \sqrt{\langle v^2 \rangle} = 512\,m/s = 512\,m/s \times 3600 = 1.84 \times 10^6\,m/hr \]

so

\[ \overline{u} = \sqrt{\langle u^2 \rangle} = 1840\,km/hr \]

\[ \overline{u} = \frac{1840\,km}{hr} \times \frac{0.621\,mi}{km} \times \frac{1\,hr}{mi} = 1143\,mi/hr \]
Review: The ideal-gas law says

\[ pV = NkT \]

where \( p \) is the pressure, \( V \) the volume, \( N \) the number of molecules, \( k \) Boltzmann's constant

\[ k = 1.38 \times 10^{-23} \text{ J/K} \]

and \( T \) the temperature in K.

Room temperature \( T_r = 295 \text{ K} \) and

\[ kT_r \approx 4.1 \text{ pN nm} = 4.1 \times 10^{-21} \text{ J} \]

Atmospheric pressure \( p_a \approx 10^5 \text{ Pa} = 10^5 \text{ N m}^{-2} \).

The mean kinetic energy of a molecule of an ideal gas at temperature \( T \) is

\[ \frac{m}{2} \langle v^2 \rangle = \frac{3}{2} kT \]

where \( m \) is the mass of the molecule.
We roughly estimate that the gravitational potential energy of an air molecule at \( z = 9 \text{ km} \) (top of Everest) is its kinetic energy:

\[
\Delta U = mgq \text{ km} \approx \frac{1}{2} m \langle v^2 \rangle.
\]

\[3J\]
So \( \frac{1}{2} \langle v^2 \rangle \approx \frac{g}{9} \text{ km} \]

\[
\langle v^2 \rangle \approx 9.8 \text{ m/s} \times 18 \times 10^3 \text{ m} = 1.76 \times 10^5 \text{ m}^2 \text{s}^{-2}.
\]

So

\[
\bar{v} = \sqrt{\langle v^2 \rangle} = \sqrt{1.76 \times 10^4} \text{ m/s} = 420 \text{ m/s}.
\]

\[3K\]
\[
\frac{3}{2} k \bar{v} \approx \frac{3}{2} 4.1 \times 10^{-3} \text{ J} = 6.15 \times 10^{-2} \text{ J} \approx \frac{1}{2} m \langle v^2 \rangle.
\]

is the mean kinetic energy of an air molecule at room temperature \( T_n = 295 \text{ K} \).

If a norm is 3 km high, then

\[
\Delta U = m \cdot g \cdot 3 \text{ km} = 4.7 \times 10^{-26} \text{ kg} \times 9.8 \text{ m/s}^2 \times 3 \text{ km} = 1.38 \times 10^{-24} \text{ J}.
\]

So \( \frac{1}{2} m \langle v^2 \rangle \) is 4500 \( \text{s}^2 \) times bigger.
The air doesn't fall to the floor because $\Delta U$ is negligible compared to $\frac{1}{2}m<v^2>$. We could make it fall by dropping the temperature to below 70K. Then the air molecules would stick together, forming a liquid on the floor.

How about a particle of dust that has the mass of a 50mm cube of water?

Now $L = 50 \times 10^{-6} \text{m} = 50 \times 10^{-4} \text{cm} \quad 50$

$V = (5 \times 10^{-3})^3 \text{cm}^3 = 125 \times 10^{-9} \text{cm}^3$

$so$

$m = 1.25 \times 10^{-7} \quad 1.25 \times 10^{-10} \text{kg}$

Now $\Delta U = mgh = 1.25 \times 10^{-10} \text{kg} \times 9.8 \text{m/s}^2 \times 3 \text{m}$

$= 3.68 \times 10^{-9} \text{J}$ which is more than $10^4$ times the available thermal energy $kT$ so the dust particle falls to the floor.
A gas of molecules escapes from a hole in the source and passes through a hole in a disk. The holes define the direction of the velocity of the molecules. Their speeds are determined when they pass through two slits in two rotating disks. The flux of molecules with the well-defined velocity \( \vec{v} \) is measured by the detector.
The data fall on the theoretical curve for all atoms and molecules of different mass \( m \) at various temperatures \( T \). This Boltzmann distribution fits all data with no adjustable parameters.

The velocity \( \bar{v} \) of any particular molecule changes as it hits other molecules, but the probability distribution \( P(\bar{v}^2) \) remains fixed at a given temperature \( T \).
The Boltzmann distribution \( P(u_x) \):

We expect that

\[ P\left( \frac{u_x}{\overline{u}_x} \right) \to 0 \text{ as } \frac{1}{u_x} \to \infty \]

We expect that

\[ P(-u_x) = P(u_x). \]

We expect that

\[ \langle u_x \rangle = 0, \text{ since no wind blows.} \]

We expect that

\[ \overline{u}_x^2 = \langle (u_x - \langle u_x \rangle)^2 \rangle = \langle u_x^2 \rangle - \langle u_x \rangle^2 \]

\[ = \langle u_x^2 \rangle - \langle u_x \rangle^2 \]

We know that

\[ \langle u_x^2 \rangle = \frac{1}{3} \langle \overline{u}^2 \rangle \text{ & } \frac{1}{2} m \langle \overline{u}^2 \rangle = \frac{3}{2} kT \]

So

\[ \overline{u}_x^2 = \langle u_x^2 \rangle = \frac{kT}{m}. \]
So now we guess that $P(v_x)$ is a Gaussian with variance $\sigma_x^2 = kT/m$.

So from p. 46 of these notes, we have

$$P(v_x) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{v_x^2}{2\sigma_x^2}}$$

$$= \frac{1}{\sqrt{\frac{kT}{m}}} \sqrt{\frac{m}{2\pi \hbar T}} e^{-\frac{m v_x^2}{2kT}}$$

We saw that for $hT \approx 4.1 \times 10^{-21}$ J, air molecules were moving at 420 m/s.

The distribution $P(v^2) = P(v_x^2)P(v_y^2)P(v_z^2)$ is

$$P(v^2) = \left(\frac{m}{2\pi \hbar T}\right)^{3/2} e^{-\frac{m}{2kT} \left(v_x^2 + v_y^2 + v_z^2\right)}$$

$$= \left(\frac{m}{2\pi \hbar T}\right)^{3/2} e^{-\frac{m v^2}{2kT}}$$
What is the probability distribution for speed \( u = \sqrt{v^2} \)? The volume element is

\[
dV = 4\pi u^2 du.
\]

Area of sphere of radius \( u \).

So since \( P(v^2) \) depends only on \( v^2 \)

\[
P(u)du = P(v^2) 4\pi u^2 du.
\]

\[
= \left( \frac{m}{2\pi\hbar T} \right)^{3/2} \exp \left( -\frac{m u^2}{2\hbar T} \right) 4\pi u^2 du.
\]

The most probable speed is the maximum of \( P(u) \).

\[
0 = \frac{dP}{du} = \left( \frac{m}{2\pi\hbar T} \right)^{3/2} \left( -\frac{m u^3}{\hbar T} + 2u \right) \exp \left( -\frac{m u^2}{2\hbar T} \right) 4\pi
\]

or

\[
\frac{-m u^3}{\hbar T} + 2u = 0 \quad \Rightarrow \quad \frac{m u^3}{\hbar T} = 2 \quad \Rightarrow \quad u^2 = \frac{2\hbar T}{m} \quad \Rightarrow \quad u = \sqrt{\frac{2\hbar T}{m}}
\]

or \( u_{mp} = 0 \times \sqrt{2} \).
The mean speed \( <u> \) is

\[
<u> = \int_0^\infty \left( \frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{m(u^2)}{2kT}} 4\pi u^2 \, du
\]

\[
= \frac{4\pi}{(\sigma \sqrt{2\pi})^3} \int_0^\infty du \, u^2 \, e^{\frac{-u^2}{2\sigma^2}}
\]

\( \sigma^2 = \sigma_x^2 = kT/m \)

Let \( y = u^2 / 2\sigma^2 \), \( dy = \frac{u}{\sigma^2} \, du \),

\[
<u> = \frac{4\pi}{(\sigma \sqrt{2\pi})^3} \int_0^\infty \sigma^2 \, dy \, 2\sigma^2 y \, e^{-y}
\]

\[
= \frac{4\pi \sigma^4}{(\sigma \sqrt{2\pi})^3} \int_0^\infty \, dy \, y \, e^{-y}
\]

0 = \int_0^\infty \int_0^\infty d(ye^{-y}) = \int_0^\infty dy \, e^{-y} - \int_0^\infty ye^{-y} \, dy \quad \text{so}

\[
\int_0^\infty dy \, ye^{-y} = \int_0^\infty dy \, e^{-y} = -e^{-y} \bigg|_0^\infty = 1.
\]
So the mean speed $\langle u \rangle$ is

$$\langle u \rangle = \frac{8 \pi \sigma}{2 \pi \sqrt{2 \pi}} = \frac{4 \sigma}{\sqrt{2 \pi}} \approx 1.596 \sigma = 1.596 \sqrt{\frac{kT}{m}}.$$  

while the most probable $u_{mp} = 5 \sqrt{2} \approx 1.414 \sigma$.

The mean speed $\langle u \rangle = 1.596 \sigma$ is slightly faster than the most probable speed $u_{mp} = 1.414 \sigma$. Why?

Because the graph $u^2 e^{-u^2/(2\sigma^2)}$ is not symmetric about its maximum $u_{mp}$, the mean speed $\langle u \rangle$ can differ from the most likely speed $u_{mp}$. Also, $\langle u \rangle$ divides the area under the curve $-u^2/(2\sigma^2)$

$$u^2 e^{-u^2/(2\sigma^2)}$$

into two equal areas, and since the curve runs to $u^{\pm \infty}$ on the right side, it's plausible that $\langle u \rangle > u_{mp}$.  

---
The probability that molecule 1 has \( \vec{v}_1 \), molecule 2 has \( \vec{v}_2 \), etc., is

\[
P(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N) = \frac{3N}{2} \left( \frac{m}{2 \pi k T} \right)^{3N/2} e^{-\frac{m \vec{v}_1^2}{2k T} - \frac{m \vec{v}_2^2}{2k T} \cdots - \frac{m \vec{v}_N^2}{2k T}} \times d^3 v_1 d^3 v_2 \cdots d^3 v_N
\]

\[
= \left( \frac{m}{2 \pi k T} \right)^{3N/2} e^{-\frac{1}{2} \cdot 2 \left( \frac{m}{k T} \right) \cdot \frac{1}{2} (\vec{v}_1^2 + \vec{v}_2^2 + \cdots + \vec{v}_N^2)} \times d^3 v_1 d^3 v_2 \cdots d^3 v_N
\]

in which \( d^3 v_i \) is a tiny cube in velocity space about \( \vec{v}_i \).

More simply, this probability distribution is

\[
P(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N) = \left( \frac{m}{2 \pi k T} \right)^{3N/2} e^{-\frac{E}{k T}}
\]

in which \( E \) is the total kinetic energy of the gas

\[
E = \sum_{i=1}^{N} \frac{1}{2m} \vec{v}_i^2
\]

If the gas is in a gravitational field \( g \), then we should include the
gravitational potential energy $mgz_i$ for each molecule at height $z_i$. (In this case,)

$$P(z_1, v_1, z_2, v_2, \ldots z_N, v_N) \propto e^{-\frac{E}{kT}}$$

(we suppress the normalization here)

where now the energy $E$

$$E = \sum_{i=1}^{N} \left( \frac{1}{2} m v_i^2 + mg z_i \right)$$

includes the gravitational energies $mgz_i$ of all the molecules.

If we include the interactions of the molecules among themselves, then we must add the interaction energy $U(x_1, x_2, \ldots x_N)$

$$P(x_1, v_1, x_2, v_2, \ldots x_N, v_N) \propto e^{-\frac{E}{kT}}$$

in which

$$E = \sum_{i=1}^{N} \frac{1}{2} m v_i^2 + U(x_1, x_2, \ldots x_N).$$

This is the Boltzmann distribution.
The probability that molecule 1 is in the tiny box of volume $d^3x$, with center $\vec{x}$, and has a velocity in the box $d^3v$, centered at $\vec{v}$, etc., is

$$d^3x_1\, d^3v_1\, d^3x_2\, d^3v_2\ldots d^3x_N\, d^3v_N\, P(\vec{x}_1, \vec{v}_1, \vec{x}_2, \vec{v}_2, \ldots \vec{x}_N, \vec{v}_N)$$

$$= N \exp \left(-\frac{E}{kT}\right)$$

where $N$ is a normalization factor, so that

$$1 = \int d^3x_1\, d^3v_1\ldots d^3x_N\, d^3v_N\, P(\vec{x}_1, \vec{v}_1, \vec{x}_2, \vec{v}_2, \ldots \vec{x}_N, \vec{v}_N).$$

and $E$ is the energy of all the molecules

$$E = \sum_{i=1}^{N} \frac{1}{2} m \vec{v}_i^2 + U(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N).$$

The main contributions to $U(\vec{x}_1, \ldots, \vec{x}_N)$ come from nearby pairs of molecules.
That is, most of $U(x_1, \ldots, x_n)$ is of the form

$$U(x'_1, \ldots, x'_n) = \sum_{i<j} U(x'_i, x'_j).$$

Note that since

$$P(\text{state}) \propto \exp \left(- \frac{E(\text{state})}{kT} \right)$$

this distribution for low temperatures

is highly peaked at the lowest energy state, which is called "the ground state."

But at high temperatures, states of high energy occur with big probabilities.

As far as we know, these formulas are exact, when suitably interpreted. We'll derive them in chapter six.
Problem 3, 2 for CO₂ & N₂

\[ P_{\text{CO}_2}(z) = A e^{-\frac{M_{\text{CO}_2} g z}{kT}} \]

\[ P_{\text{N}_2}(z) = B e^{-\frac{M_{\text{N}_2} g z}{kT}} \]

where \( \frac{A}{B} = \frac{0.0335}{78.084} = 0.000429 \approx 4.3 \times 10^{-4} \)

\[ M_{\text{N}_2} = \frac{28 g}{N_A} = \frac{28 \times 10^{-3} kg}{6 \times 10^{23}} = 4.7 \times 10^{-26} kg \]

\[ M_{\text{CO}_2} = \frac{12 + 32 g}{N_A} = \frac{44 \times 10^{-3} kg}{6 \times 10^{23}} = 7.3 \times 10^{-26} kg \]

So

\[ \frac{P_{\text{CO}_2}(z)}{P_{\text{N}_2}(z)} = \frac{A}{B} \]

\[ -(M_{\text{CO}_2} - M_{\text{N}_2}) g z / (kT) \]

is the ratio at height \( z \).
With $g = 10 \text{ m} \text{s}^{-2}$ and $z = 10 \text{ km} = 10^4 \text{ m}$, and $kT_o = 4.1 \times 10^{-21} \text{ J}$, we have

$$\frac{P_{co_2}(10)}{P_{N_2}(10)} = 4.3 \times 10^{-4} \exp \left( -\frac{(7.3 - 4.7)10 \times 10^{-5}}{4.1 \times 10^{-21} \text{ J}} \right)$$

$$= 4.3 \times 10^{-4} \times e^{-2.6/4.1}$$

$$= 4.3 \times 10^{-4} \times e^{-0.634}$$

$$= 4.3 \times 10^{-4} \times 0.53$$

$$= 2.28 \times 10^{-4}$$

So, the carbon-dioxide level at 10 km is only 53% of what it is at sea level.
Osmotic Pressure

Suppose the membrane passes water but not some other solute, for instance, sugar.

Then the concentration $c_\text{w} = \frac{N_w}{V}$ will be the same on both sides of the membrane, but if we put sugar to the right of the membrane, then $c_s = \frac{N_s}{V}$ would be zero on the left and positive on the right.