

5

Continuous Random Numbers

It may be that the race is not always to the swift nor the battle to the strong—but that's the way to bet.

Damon Runyon

In Chapter 3 we discussed the basics of probability theory for discrete events. However, the components of the momenta of the particles in an ideal gas are continuous variables. This requires an extension of probability theory to deal with continuous random numbers, which we will develop in this chapter.

5.1 Continuous Dice and Probability Densities

To illustrate the essential features of continuous random numbers we will use a simple generalization of throwing a die. While a real die has six distinct sides, our continuous die can take on any real value, x , between 0 and 6.

Assume that our continuous die is ‘honest’ in the sense that all real numbers between 0 and 6 are equally probable. Since there is an infinite number of possibilities, the probability of any particular number being found is $p = 1/\infty = 0$.

On the other hand, if we ask for the probability of the continuous random number x being in some interval, the probability can be non-zero. If all values of x are equally probable, we would assume that the probability of finding x in the interval $[a, b]$ is proportional to the length of that interval,

$$P([a, b]) = A \int_a^b dx = A(b - a) \quad (5.1)$$

where A is a normalization constant. Since the probability of x being somewhere in the interval $[0, 6]$ must be 1,

$$P([0, 6]) = A \int_0^6 dx = 6A = 1, \quad (5.2)$$

we have $A = 1/6$.

This leads us to define a ‘probability density function’ or simply a ‘probability density’, $P(x) = 1/6$, to provide us with an easy way of computing the probability of finding x in the interval $[a, b]$,

$$P([a, b]) = \int_a^b P(x) dx = \int_a^b \frac{1}{6} dx = \frac{b-a}{6}. \quad (5.3)$$

I apologize for using P to denote both probabilities and probability densities, even though they are very different things. This confusing notation certainly does not help keep the distinction between the two concepts in mind. On the other hand, it is usually easy to tell which concept is intended. Since much of the literature assumes that the reader can figure it out, this is probably a good place to become accustomed to doing so.

5.2 Probability Densities

We can use the idea of a probability density to extend consideration to cases in which all values of a continuous random number are not equally likely. The only change is that $P(x)$ is then no longer a constant. The probability of finding x in an interval is given by the integral over that interval,

$$P([a, b]) = \int_a^b P(x) dx. \quad (5.4)$$

Generalization to multiple dimensions is simply a matter of defining the probability density of a multidimensional function. If there are two continuous random numbers, x and y , the probability density is $P(x, y)$.

There is one feature of probability densities that might not be clear from our example of continuous dice. In general, although probabilities are dimensionless, probability densities have units. For example, if the probability of finding a single classical particle is uniform in a box of volume V , the probability density within the volume is

$$P(x, y, z) = \frac{1}{V}, \quad (5.5)$$

and it has units of $[m^{-3}]$. When the probability density is integrated over a sub-volume of the box, the result is a dimensionless probability, as expected.

Because probability densities have units, there is no upper limit on their values. Their values must, of course, be positive. However, unlike probabilities, they do not have to be less than 1. In fact, it is even possible for a probability density to diverge at one or more points, as long as the integral over all values is 1.

Like probabilities, probability densities must be normalized. If Ω indicates the entire range over which the probability density is defined, then

$$\int_{\Omega} P(x) dx = 1 \quad (5.6)$$

with obvious extensions to multi-dimensional, continuous random numbers.

Marginal probabilities are defined in analogy to the definition for discrete numbers

$$P_x(x) = \int_{-\infty}^{\infty} P(x, y) dy. \quad (5.7)$$

Conditional probabilities can also be defined if $P_x(x) \neq 0$. The conditional probability $P(y|x)$ can be written as

$$P(y|x) = \frac{P(x, y)}{P_x(x)}. \quad (5.8)$$

In general, we have

$$\begin{aligned} P(x, y) &= P(x|y)P_y(y) \\ &= P(y|x)P_x(x). \end{aligned} \quad (5.9)$$

Bayes' theorem for continuous variables then takes the form

$$P(y|x) = \frac{P(x|y)P_y(y)}{P_x(x)}. \quad (5.10)$$

Independence is also defined in analogy to discrete probability theory. Two continuous random numbers are said to be independent if

$$P(x, y) = P_x(x)P_y(y). \quad (5.11)$$

Using eq. (5.9), we can see that if x and y are independent random numbers and $P_y(y) \neq 0$, the conditional probability $P(x|y)$ is independent of y :

$$P(x|y) = \frac{P(x, y)}{P_y(y)} = \frac{P_x(x)P_y(y)}{P_y(y)} = P_x(x). \quad (5.12)$$

The average of any function $F(x)$ can be calculated by integrating over the probability density

$$\langle F(x) \rangle = \int_{-\infty}^{\infty} F(x)P(x)dx. \quad (5.13)$$

Averages, moments, and central moments are all defined as expected from discrete probability theory.

Mean:

$$\langle x \rangle = \int_{-\infty}^{\infty} xP(x)dx \quad (5.14)$$

n -th moment:

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n P(x)dx \quad (5.15)$$

n -th central moment:

$$\langle (x - \langle x \rangle)^n \rangle = \int_{-\infty}^{\infty} (x - \langle x \rangle)^n P(x)dx. \quad (5.16)$$

As for discrete probabilities, the variance is defined as the second central moment, and the standard deviation is the square root of the variance.

A word of warning:

Many books refer to continuous random numbers as if they were discrete. They will refer to the probability of finding a random number in a given interval as being proportional to the ‘number of points’ in that interval. This is entirely incorrect and can cause considerable confusion. The number of points in an interval is infinite. It is even the same infinity as the number of points in any other interval, in the sense that the points in any two intervals can be mapped one-to-one onto each other. This error seems to have arisen in the nineteenth century, before physicists were entirely comfortable working with continuous random numbers. Although we are now in the twenty-first century, the error seems to be quite persistent. I can only hope that the next generation of physicists will finally overcome it.

5.3 Dirac Delta Functions

In Chapter 3 we discussed finding the probability of the sum of two dice using Kronecker delta functions. In the next section we will carry out the analogous calculation for two continuous dice. Although the problem of two continuous dice is highly artificial, it is useful because its mathematical structure will often be encountered in statistical mechanics.

For these calculations we will introduce the Dirac delta function, named for its inventor, Paul Adrien Maurice Dirac (English physicist, 1902–1984). The Dirac delta

function is an extension of the idea of a Kronecker delta to continuous functions. This approach is not usually found in textbooks on probability theory—which is unfortunate, because Dirac delta functions make calculations much easier.

Dirac delta functions are widely used in quantum mechanics, so most physics students will have encountered them before reading this book. However, since we will need more properties of the delta function than are usually covered in courses in quantum mechanics, we will present a self-contained discussion in the next section.

5.3.1 Definition of Delta Functions

To provide a simple definition of a Dirac delta function, we will first consider an ordinary function that is non-zero only in the neighborhood of the origin,¹

$$\delta_\epsilon(x) \equiv \begin{cases} 0 & x < -\epsilon \\ \frac{1}{2\epsilon} & -\epsilon \leq x \leq \epsilon \\ 0 & x > \epsilon \end{cases}.$$

Clearly, this function is normalized,

$$\int_{-\infty}^{\infty} \delta_\epsilon(x) dx = 1. \quad (5.17)$$

In fact, since $\delta_\epsilon(x)$ is only non-zero close to the origin,

$$\int_a^b \delta_\epsilon(x) dx = 1 \quad (5.18)$$

as long as $a < -\epsilon$ and $b > \epsilon$.

It is also obvious that the function is symmetric,

$$\delta_\epsilon(x) = \delta_\epsilon(-x). \quad (5.19)$$

If we consider the function $\delta_\epsilon(cx)$, where c is a constant, the symmetry of the function means that the sign of c is irrelevant,

$$\delta_\epsilon(cx) = \delta_\epsilon(|c|x). \quad (5.20)$$

Note that the width of the function $\delta_\epsilon(cx)$ is a factor of $1/|c|$ times that of $\delta_\epsilon(x)$, while the height of the function remains the same ($1/2\epsilon$).

¹ An alternative way of defining the Dirac delta function uses Gaussian functions, which has advantages when discussing derivatives. The approach used in this section has been chosen for its simplicity.

5.3.2 Integrals over Delta Functions

We can integrate over $\delta_\epsilon(cx)$ by defining a new variable $y = cx$,

$$\int_a^b \delta_\epsilon(cx) dx = \int_{a/|c|}^{b/|c|} \delta_\epsilon(y) dy / |c| = \frac{1}{|c|}. \quad (5.21)$$

When the argument of the delta function has a zero at some value of $x \neq 0$, the integral is unaffected as long as the limits of the integral include the zero of the argument of the delta function. As long as $a < d/c - \epsilon$ and $b > d/c + \epsilon$, we have

$$\int_a^b \delta_\epsilon(cx - d) dx = \int_{a/|c|}^{b/|c|} \delta_\epsilon(y) dy / |c| = \frac{1}{|c|}. \quad (5.22)$$

The Dirac delta function, $\delta(x)$, is defined as the limit of $\delta_\epsilon(x)$ as ϵ goes to zero.

$$\delta(x) \equiv \lim_{\epsilon \rightarrow 0} \delta_\epsilon(x). \quad (5.23)$$

The Dirac delta function is zero for all $x \neq 0$ and infinite for $x = 0$, so that calling it a function greatly annoys most mathematicians. For this reason, mathematics books on probability theory rarely mention the delta function. Nevertheless, if you are willing to put up with a mathematician's disapproval, the Dirac delta function can make solving problems much easier.

The integral of the Dirac delta function is also unity when the limits of integration include the location of the zero of the argument of the delta function,

$$\int_a^b \delta(x) dx = 1 \quad (5.24)$$

as long as $a < 0$ and $b > 0$, and zero otherwise.

Similarly, when the argument of the delta function is $(cx - d)$, the value of the integral is

$$\int_a^b \delta(cx - d) dx = \frac{1}{|c|} \quad (5.25)$$

as long as $a < d/c$ and $b > d/c$, and zero otherwise.

5.3.3 Integral of $f(x)$ times a Delta Function

The Dirac delta function can be used to pick out the value of a continuous function at a particular point, just as the Kronecker delta does for a discrete function. This is one of its most useful properties.

Consider a function $f(x)$ that is analytic in some region that includes the point x_0 . Then we can expand the function as a power series with a non-zero radius of convergence

$$f(x) = \sum_{j=0}^{\infty} \frac{1}{j!} f^{(j)}(x_0) (x - x_0)^j, \quad (5.26)$$

where $f^{(j)}(x_0)$ is the j -th derivative of $f(x)$, evaluated at $x = x_0$,

$$f^{(j)}(x_0) = \left. \frac{d^j}{dx^j} f(x) \right|_{x=x_0}. \quad (5.27)$$

Consider the integral over the product of $f(x)$ and $\delta_\epsilon(x - x_0)$

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta_\epsilon(x - x_0) dx &= \frac{1}{2\epsilon} \int_{x_0-\epsilon}^{x_0+\epsilon} f(x) dx \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} \frac{1}{2\epsilon} \int_{x_0-\epsilon}^{x_0+\epsilon} f^{(j)}(x_0) (x - x_0)^j dx \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} \frac{1}{2\epsilon} f^{(j)}(x_0) \frac{1}{j+1} \left[\epsilon^{j+1} - (-\epsilon)^{j+1} \right]. \end{aligned} \quad (5.28)$$

All terms with odd j vanish. Keeping only the even terms and defining $n = j/2$ for even values of j , we can write the integral as

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \delta_\epsilon(x - x_0) dx &= \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} f^{(2n)}(x_0) \epsilon^{2n} \\ &= f(x_0) + \sum_{n=1}^{\infty} \frac{1}{(2n+1)!} f^{(2n)}(x_0) \epsilon^{2n}. \end{aligned} \quad (5.29)$$

When we take the limit of $\epsilon \rightarrow 0$, the function, $\delta_\epsilon(\cdot)$ becomes a Dirac delta function, and the right-hand side of the equation is just $f(x_0)$,

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0). \quad (5.30)$$

We can generalize eq. (5.30) to include more general arguments of the delta function. First, if the argument is of the form $c(x - x_0)$, the result is divided by $|c|$,

$$\int_{-\infty}^{\infty} f(x) \delta(c(x - x_0)) dx = \frac{f(x_0)}{|c|}. \quad (5.31)$$

We can further generalize this equation to allow the delta function to have an arbitrary argument, $g(x)$. If $g(x)$ has zeros at the points $\{x_j | j = 1, \dots, n\}$, and $g'(x)$ is the derivative of $g(x)$, the integral becomes

$$\int_{-\infty}^{\infty} f(x) \delta(g(x)) dx = \sum_{j=1}^n \frac{f(x_j)}{|g'(x_j)|}. \quad (5.32)$$

This can be proved by expanding $g(x)$ about each of its zeros.

Common errors in working with Dirac delta functions:

- Not including all zeros of $g(x)$ inside the region of integration.
- Including zeros of $g(x)$ that are outside the region of integration.
- Not using the absolute value of $g'(x)$ in the denominator.

5.4 Transformations of Continuous Random Variables

We are often interested in functions of random variables, which are themselves random variables. As an example, consider the following problem.

Given two continuous random variables, x and y , along with their joint probability density, $P(x, y)$, we wish to find the probability density of a new random variable, s , that is some function of the original random variables, $s = f(x, y)$. The formal solution can be written in terms of a Dirac delta function,

$$P(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) \delta(s - f(x, y)) dx dy. \quad (5.33)$$

As was the case for the corresponding discrete random variables, the probability density of s is automatically normalized,

$$\begin{aligned} \int_{-\infty}^{\infty} P(s) ds &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) \delta(s - f(x, y)) ds dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) dx dy = 1. \end{aligned} \quad (5.34)$$

The last equality is due to the normalization of $P(x, y)$.

5.4.1 Sum of Two Random Numbers

To see how eq. (5.33) works in practice, consider the sum of two uniformly distributed random numbers. We will use the example of two continuous dice, in which the two

random numbers x and y each take on values in the interval $[0, 6]$, and the probability density is uniform,

$$P(x, y) = 1/36. \quad (5.35)$$

This probability density is consistent with the normalization condition in eq. (5.6).

The probability density for $s = x + y$ is then

$$\begin{aligned} P(s) &= \int_0^6 \int_0^6 P(x, y) \delta(s - (x + y)) dx dy \\ &= \frac{1}{36} \int_0^6 \int_0^6 \delta(s - (x + y)) dx dy. \end{aligned} \quad (5.36)$$

Following the same procedure as in Section 3.4, we first carry out the integral over y .

$$\int_0^6 \delta(s - (x + y)) dy = \begin{cases} 1 & 0 \leq s - x \leq 6 \\ 0 & \text{otherwise} \end{cases}. \quad (5.37)$$

In direct analogy to the procedure for calculating the sum of two discrete dice in Section 3.4, we have two conditions on the remaining integral over x . Only those values of x for which both $x < s$ and $x > s - 6$ contribute to the final answer. These limits are in addition to the limits of $x < 6$ and $x > 0$ that are already explicit in the integral over x . Since all four of these inequalities must be satisfied, we must take the more restrictive of the inequalities in each case. Which inequality is the more restrictive depends on the value of s . Determining the proper limits on the integral over x is illustrated in Table 5.1, which should be compared to Table 3.3 for the corresponding discrete case.

For $s \leq 6$, the lower bound on x is 0 and the upper bound is s . For $s \geq 6$, the lower bound on x is $s - 6$ and the upper bound is 6. The sums can then be evaluated explicitly,

Table 5.1 *Determining the limits for the second integral when evaluating eq. (5.36).*

	lower limit	upper limit
From limits on integral:	$x \geq 0$	$x \leq 6$
From delta function:	$x \geq s - 6$	$x \leq s$
More restrictive if $s \leq 6$	$x \geq 0$	$x \leq s$
More restrictive if $s \geq 6$	$x \geq s - 6$	$x \leq 6$

$$P(s) = \begin{cases} \int_0^s \frac{1}{36} = \frac{s}{36} & s \leq 6 \\ \int_{s-6}^6 \frac{1}{36} = \frac{12-s}{36} & s \geq 6 \end{cases}. \quad (5.38)$$

5.5 Bayes' Theorem

Bayes' theorem, eq. (3.13), was derived in Section 3.3.1,

$$P(A|B) = \frac{P(B|A)P_A(A)}{P_B(B)}, \quad (5.39)$$

and given in eq. (5.10) for continuous variables as

$$P(y|x) = \frac{P(x|y)P_y(y)}{P_x(x)}. \quad (5.40)$$

If we accept the Bayesian definition of probability in Section 3.1, we can apply eq. (5.39) to the determination of theoretical parameters from experiment.

Let X denote the data from an experiment, while θ denotes the theoretical parameter(s) we wish to determine. Using standard methods, we can calculate the conditional probability $P(X|\theta)$ of observing a particular set of data if we know the parameters. However, that is not what we want. We have the results of the experiment and we want to calculate the parameters conditional on that data.

Bayes' theorem gives us the information we want about the theoretical parameters after the experiment. It provides the information on a particular set of measurements X , as the conditional probability for θ in the following equation:

$$P(\theta|X) = \frac{P(X|\theta)P_\theta(\theta)}{P_X(X)}. \quad (5.41)$$

In Bayesian terminology, $P_\theta(\theta)$ is called the 'prior' and represents whatever knowledge we had of the parameters before we carried out the experiment. The conditional probability $P(X|\theta)$ —viewed as a function of θ —is called the 'likelihood'. The conditional probability $P(\theta|X)$ is known as the 'posterior', and represents our knowledge of the parameters after we have obtained data from an experiment.

For example, suppose we want to determine the value of the asymptotic frequency f (the frequentist's probability) for a particular random event. Assume we have done an experiment with a binary outcome of success or failure, and we obtained $n = 341,557$ successes in $N = 10^6$ trials. The experimental data, denoted previously by X , is n

in this case. We would immediately guess that $f \approx 0.341557$, but how close to that value is it?

Our theoretical parameter, denoted by θ in eq. (5.41), is the value of f . Suppose we knew nothing of the value of f before the experiment, except that $0 \leq f \leq 1$. A reasonable description of our knowledge (or lack of it) would be to say that our prior is a uniform constant for values between zero and 1. From the normalization condition, the constant must be 1,

$$P_f(f) = \begin{cases} 0 & f < 0 \\ 1 & 0 \leq f \leq 1 \\ 0 & f > 1 \end{cases} \quad (5.42)$$

The likelihood is given by the binomial distribution

$$P_N(n|f) = \frac{N!}{n!(N-n)!} f^n (1-f)^{N-n}. \quad (5.43)$$

As is usual in Bayesian calculations, we will ignore $P_{N,n}(n)$; since it does not depend on f , it simply plays the role of a normalization constant. Putting eqs. (5.41), (5.42), and (5.43) together and ignoring multiplicative constants, we have the probability density for f ,

$$P_N(f|n) \propto f^n (1-f)^{N-n}. \quad (5.44)$$

For large n and N , this is a sharply peaked function. The maximum can be found by setting the derivative of its logarithm equal to zero:

$$\frac{d}{df} P_N(f|n) = \frac{d}{df} [n \ln f + (N-n) \ln(1-f) + \text{constants}] = \frac{n}{f} - \frac{N-n}{1-f} = 0. \quad (5.45)$$

The maximum is therefore located at

$$f_{\max} = \frac{n}{N} \quad (5.46)$$

as expected, since we had n successes in N trials.

The variance can be found by approximating the distribution by a Gaussian, finding the second derivative, and evaluating it at f_{\max} ,

$$\frac{d^2}{df^2} [n \ln f + (N-n) \ln(1-f)] = \frac{d}{df} \left[\frac{n}{f} - \frac{N-n}{1-f} \right] \quad (5.47)$$

$$= -\frac{n}{f^2} - \frac{N-n}{(1-f)^2}. \quad (5.48)$$

Evaluating this derivative at $f = f_{\max} = n/N$ The variance is then given by

$$-\frac{1}{\sigma^2} = -\frac{N}{f_{\max}} - \frac{N}{1-f_{\max}} = -\frac{N}{f_{\max}(1-f_{\max})} \quad (5.49)$$

or

$$\sigma^2 = \frac{f_{\max}(1-f_{\max})}{N}. \quad (5.50)$$

The standard deviation is then

$$\sigma = \sqrt{\frac{f_{\max}(1-f_{\max})}{N}} \quad (5.51)$$

so that the uncertainty in the value of f is proportional to $1/\sqrt{N}$. In the case of our example with $f \approx 0.341557$ and $N = 10^6$, the standard deviation is ≈ 0.00047 .

5.6 Problems

PROBLEM 5.1

Probability densities

1. Given a continuous random number x , with the probability density

$$P(x) = A \exp(-2x)$$

for all $x > 0$, find the value of A and the probability that $x > 1$.

2. It is claimed that the function

$$P(y) = \frac{B}{y}$$

where B is a constant, is a valid probability density for $0 < y < 1$.

Do you agree? If so, what is the value of B ?

3. It is claimed that the function

$$P(y) = \sqrt{\frac{C}{y}}$$

where C is a constant, is a valid probability distribution for $0 < y < 1$.

Do you agree? If so, what is the value of C ?

4. The probability distribution $P(x, y)$ has the form

$$P(x, y) = Dxy$$

for the two random numbers x and y in the region $x > 0, y > 0$, and $x + y < 1$, where D is a constant, and $P(x, y) = 0$ outside this range.

- What is the value of the constant D ?
- What is the probability that $x < 1/2$?
- Are x and y independent?

PROBLEM 5.2

How well can we measure the asymptotic frequency?

I have run a number of computer simulations in which a ‘success’ occurred with probability p . It would be no fun if I told you the value of p , but I will tell you the results of the simulations.

<i>Trials</i>	<i>Successes</i>
10^2	69
10^3	641
10^4	6353
10^5	63738
10^6	637102
10^7	6366524

- Using Bayesian statistics, we showed that we can represent our knowledge of the value of p by a probability distribution (or probability density) $P(p|n)$, which is the conditional probability density for p , given that an experiment of N trials produced n successes. $P(p|n)$ is found using Bayes’ theorem. Note that $P(n|p)$, the conditional probability of finding n successes in N trials when the probability is p , is just the binomial distribution that we have been studying.

Find the location of the maximum of this probability density by taking the first derivative of $P(p|n)$ and setting

$$\frac{\partial}{\partial p} \ln P(p|n) = 0.$$

Find the general expression for the width of $P(p|n)$ by taking the second derivative. The procedure is essentially the same as the one we used for calculating the Gaussian approximation to the integrand in deriving Stirling’s approximation.

- Using the result you obtained above, what can you learn about the value of p from each of these trials? Is the information from the various trials consistent? (Suggestion: A spreadsheet can be very helpful for calculations in this problem.)

PROBLEM 5.3

Continuous probability distributions

Determine the normalization constant and calculate the mean, variance, and standard deviation for the following random variables:

1. For $x \geq 0$

$$P(x) = A \exp(-ax).$$

2. For $x \geq 1$

$$P(x) = Bx^{-3}.$$

PROBLEM 5.4

Integrals over delta functions

Evaluate the following integrals in closed form:

- 1.

$$\int_{-\infty}^{\infty} x^4 \delta(x^2 - y^2) dx$$

- 2.

$$\int_{-1}^{\infty} \exp(-x) \delta(\sin(x)) dx.$$

3. The probability density $P(x, y)$ has the form

$$P(x, y) = 24xy$$

for the two random numbers x and y in the range $x > 0$, $y > 0$, and $x + y < 1$, and $P(x, y) = 0$ outside this range.

What is the probability distribution for the following new random variable,

$$z = x + y?$$

PROBLEM 5.5

Transforming random variables

Consider the two random variables x and y . Their joint probability density is given by

$$P(x, y) = A \exp[-x^2 - 2y]$$

where A is a constant, $x \geq 0$, $y \geq 0$, and $y \leq 1$.

1. Evaluate the constant A .
2. Are x and y independent? Justify your answer.
3. If we define a third random number by the equation

$$z = x^2 + 2y$$

what is the probability density $P(z)$?

PROBLEM 5.6

Maxwell–Boltzmann distribution

In the near future we will derive the Maxwell–Boltzmann distribution for the velocities (or the momenta) of gas particles. The probability density for the x -component of the velocity is

$$P(v_x) = A \exp \left[-\beta \frac{1}{2} m v_x^2 \right].$$

1. What is the value of the normalization constant A ?
2. What is the probability density for the velocity $\vec{v} = (v_x, v_y, v_z)$?
3. What is the probability distribution (probability density) of the speed (magnitude of the velocity)?

PROBLEM 5.7

Two one-dimensional, ideal-gas particles

Consider two ideal-gas particles with masses m_A and m_B , confined to a one-dimensional box of length L . Assume that the particles are in thermal equilibrium with each other, and that the total kinetic energy is $E = E_A + E_B$. Use the usual assumption that the probability is uniform in phase space, subject to the constraints.

1. Calculate the probability distribution $P(E_A)$ for the energy of one of the particles.
2. Calculate the average energy of the particle, $\langle E_A \rangle$.
3. Find the most probable value of the energy E_A (the location of the maximum of the probability density).

PROBLEM 5.8

Energy distribution of a free particle

Suppose we have an ideal gas in a finite cubic box with sides of length L . The system is in contact with a thermal reservoir at temperature T . We have determined the momentum distribution of a single particle to be:

$$P(\vec{p}) = X \exp \left[-\beta \frac{|\vec{p}|^2}{2m} \right].$$

Find $P(E)$, where E is the energy of a single particle.

PROBLEM 5.9

Particle in a gravitational field

An ideal gas particle in a gravitational field has a probability distribution in momentum \vec{p} and height z of the form

$$P(\vec{p}, z) = X \exp \left[-\beta \frac{|\vec{p}|^2}{2m} - \beta mgz \right]$$

for $0 \leq z < \infty$ and all values of momentum.

1. Evaluate the constant X .
2. Calculate the average value of the height: $\langle z \rangle$.
3. Calculate the probability distribution of the total energy of the particle

$$E = \frac{|\vec{p}|^2}{2m} + mgz.$$
