Appendix B
Probabilities and distributions

B.1 Expectation value and variance

Definition B.1. Suppose a (not necessarily quantum) experiment to measure a quantity Q can yield any one of N possible outcomes $\{Q_i\}$ $(1 \le i \le N)\}$, with respective probabilities pr_i. Then Q is called a *random variable* and the set of values $\{\mathrm{pr}_i\}$ for all values of i is called the *probability distribution*. The *expectation (mean) value* of Q is

$$\langle Q \rangle = \sum_{i=1}^{N} \mathrm{pr}_i Q_i. \tag{B.1}$$

Exercise B.1. Find the expectation of the value displayed on the top face of a fair die.

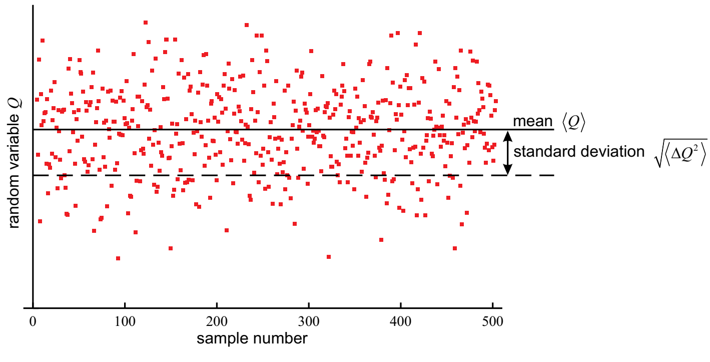

Fig. B.1 Mean and rms standard deviation of a random variable.

Definition B.2. The *mean square variance* of random variable Q is

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1

$$\left\langle \Delta Q^2 \right\rangle = \left\langle (Q - \langle Q \rangle)^2 \right\rangle = \sum_i \mathrm{pr}_i \left(Q_i - \langle Q \rangle\right)^2 . \tag{B.2}$$

The *root mean square (rms) standard deviation*, or *uncertainty* of Q is then $\sqrt{\langle \Delta Q^2 \rangle}$.

While the expectation value, $\langle Q \rangle = \sum_{i=1}^{N} \mathrm{pr}_i Q_i$, shows the mean measurement output, the statistical uncertainty shows by how much, on average, a particular measurement result will deviate from the mean (Fig. B.1).

Exercise B.2. Show that, for any random variable Q,

$$\left\langle \Delta Q^2 \right\rangle = \left\langle Q^2 \right\rangle - \langle Q \rangle^2 . \tag{B.3}$$

Exercise B.3. Calculate the mean square variance of the value displayed on the top face of a fair die. Show by direct calculation that Eqs. (B.2) and (B.3) yield the same.

Exercise B.4. Two random variables Q and R are independent, i.e., the realization of one does not affect the probability distribution of the other (for example, a die and a coin being tossed next to each other). Show that $\langle QR \rangle = \langle Q \rangle \langle R \rangle$. Is this statement valid if Q and R are not independent?
Hint: Independence means that events Q_i and R_j occur at the same time with probability $\mathrm{pr}_i^Q \mathrm{pr}_j^R$ for each pair (i, j), where pr_i^Q is the probability of the ith value of variable Q and pr_j^R is the probability of the jth value of R.

Exercise B.5. Suppose a random variable Q is measured (for example, a die is thrown) N times. Consider the random variable $\tilde{Q}$ that is the sum of the N outcomes. Show that the expectation and variance of $\tilde{Q}$ equal

$$\left\langle \tilde{Q} \right\rangle = N \langle Q \rangle$$

and

$$\left\langle \Delta \tilde{Q}^2 \right\rangle = N \left\langle \Delta Q^2 \right\rangle ,$$

respectively.

B.2 Conditional probabilities

The *conditional probability* $\mathrm{pr}_{A|B}$ is the probability of some event A given that another event, B, is known to have occurred. Examples are:

- the probability that the value on a die is odd given that it is greater than 3;
- the probability that Alice's HIV test result will be positive given that she is actually not infected;
- the probability that Bob plays basketball given that he is a man and 185 cm tall;

- the probability that it will rain tomorrow given that it has rained today.

Let us calculate the conditional probability using the third example. Event A is "Bob plays basketball". Event B is "Bob is a 185-cm tall man". The conditional probability is equal to the number $N(A \text{ and } B)$ of 185-cm tall men who play basketball divided by the number $N(B)$ of 185-cm tall men [Fig. B.2(a)]:

$$\mathrm{pr}_{A|B} = \frac{N(A \text{ and } B)}{N(B)}. \tag{B.4}$$

Let us divide both the numerator and the denominator of the above fraction by N, the total number of people in town. Then we have in the numerator $N(A \text{ and } B)/N = \mathrm{pr}_{A \text{ and } B}$ — the probability that a randomly chosen person is a 185-cm tall man who plays basketball, and in the denominator, $N(B)/N = \mathrm{pr}_B$ — the probability that a random person is a 185-cm tall man. Hence

$$\mathrm{pr}_{A|B} = \frac{\mathrm{pr}_{A \text{ and } B}}{\mathrm{pr}_B}. \tag{B.5}$$

This is a general formula for calculating conditional probabilities.

Exercise B.6. Suppose events $B_1, \ldots, B_n$ are mutually exclusive and collectively exhaustive, i.e., one of them must occur, but no two occur at the same time [Fig. B.2(b)]. Show that, for any other event A,

$$\mathrm{pr}_A = \sum_{i=1}^{n} \mathrm{pr}_{A|B_i} \mathrm{pr}_{B_i}. \tag{B.6}$$

This result is known as the *theorem of total probability*.

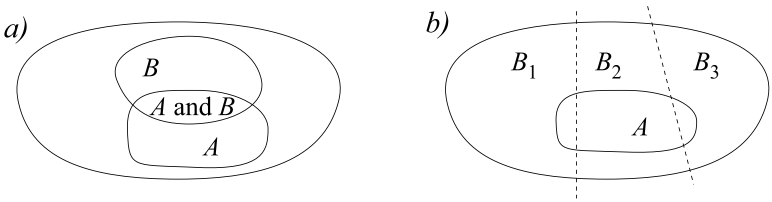

Fig. B.2 Conditional probabilities. a) Relation between the conditional and combined probabilities, Eq. (B.5). b) Theorem of total probability (Ex. B.6).

Exercise B.7. The probability that a certain HIV test gives a false positive result is

$$\mathrm{pr}_{\text{positive}|\text{not infected}} = 0.05.$$

The probability of a false negative result is zero. It is known that, of all people taking the test, the probability of actually being infected is $\mathrm{pr}_{\text{infected}} = 0.001$.

a) What is the probability $\mathrm{pr}_{\text{positive and not infected}}$ that a random person taking the test is not infected *and* shows a false positive result?
b) What is the probability $\mathrm{pr}_{\text{positive}}$ that a random person taking the test shows a positive result?
c) A random person, Alice, has been selected and the test has been performed on her. Her result turned out to be positive. What is the probability that she is not infected?

Hint: To visualize this problem, imagine a city of one million. How many of them are infected? How many are not? How many positive test results will there be all together?

B.3 Binomial and Poisson distributions

Exercise B.8. A coin is tossed n times. Find the probability that heads will appear k times, and tails $n-k$ times:

a) for a fair coin, i.e., the probability of getting heads or tails in a single toss is $1/2$;
b) for a *biased* coin, with the probabilities for the heads and tails being p and $1-p$, respectively.
Answer:
$$\mathrm{pr}_k = \binom{n}{k} p^k (1-p)^{n-k}. \tag{B.7}$$

The probability distribution defined by Eq. (B.7) is called the *binomial* distribution. We encounter this distribution in everyday life, often without realizing it. Here are a few examples.

Exercise B.9.§

a) On a given day in a certain city 20 babies were born. What is the probability that exactly nine of them are girls?
b) A student answers $3/4$ of questions on average. What is the probability that (s)he scores perfectly on a 10-question test?
c) A certain politician has 60% electoral support. What is the probability that (s)he receives more than 50% of the votes in a district with 100 voters?

Exercise B.10. Find the expectation value and the uncertainty of the binomial distribution (B.7).
Answer:
$$\langle k \rangle = np; \quad \langle \Delta k^2 \rangle = np(1-p). \tag{B.8}$$

Exercise B.11. In a certain big city, 10 babies are born per day on average. What is the probability that on a given day, exactly 12 babies are born?

a) The city population is 100000.

b) The city population is 1000000.

Hint: Perhaps there is a way to avoid calculating 1000000!.

We see from the above exercise that in the limit $p \to 0$ and $n \to \infty$, but $\lambda = pn =$ const, the probabilities in the binomial distribution become dependent on λ, rather than p and n individually. This important extension of the binomial distribution is known as the *Poisson (Poissonian)* distribution.

Exercise B.12. Show that in the limit $p \to 0$ and $n \to \infty$, but $\lambda = pn =$ const, the binomial distribution (B.7) becomes

$$\mathrm{pr}_k = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \tag{B.9}$$

using the following steps.

a) Show that $\lim_{n\to\infty} \frac{1}{n^k} \binom{n}{k} = \frac{1}{k!}$.
b) Show that $\lim_{n\to\infty} (1-p)^{n-k} = \mathrm{e}^{-\lambda}$.
c) Obtain Eq. (B.9).

Exercise B.13. Find the answer to Ex. B.11 in the limit of an infinitely large city.

Here are some more examples of the Poisson distribution.

Exercise B.14.[§]

a) A patrol policeman posted on a highway late at night has discovered that, on average, 60 cars pass every hour. What is the probability that, within a given minute, exactly one car will pass that policeman?
b) A cosmic ray detector registers 500 events per second on average. What is the probability that this number equals exactly 500 within a given second?
c) The average number of lions seen on a one-day safari is 3. What is the probability that, if you go on that safari, you will not see a single lion?

Exercise B.15. Show that both the mean and variance of the Poisson distribution (B.9) equal λ.

For example, in a certain city, 25 babies are born per day on average, so $\lambda = 25$. The root mean square uncertainty in this number $\sqrt{\lambda} = 5$, i.e., on a typical day we are much more likely to see 20 or 30 babies rather than 10 or 40 (Fig. B.3).

Although the absolute uncertainty of n increases with $\langle n \rangle$, the *relative* uncertainty $\sqrt{\lambda}/\lambda$ decreases. In our example above, the relative uncertainty is $5/25 = 20\%$. But in a smaller city, where $\langle n \rangle = 4$, the relative uncertainty is as high as $2/4 = 50\%$.

B.4 Probability densities

So far, we have studied random variables that can take values from a discrete set, with the probability of each value being finite. But what if we are dealing with

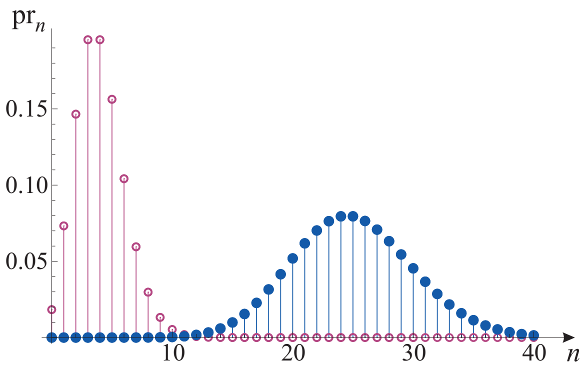

Fig. B.3 the Poisson distribution with $\langle n \rangle = 4$ (empty circles) and $\langle n \rangle = 25$ (filled circles).

a *continuous* random variable — for example, the wind speed, the decay time of a radioactive atom, or the range of a projectile? In this case, there is now way to assign a finite probability value to each specific value of Q. The probability that the atom decays after *precisely* two milliseconds, or the wind speed is *precisely* five meters per second, is infinitely small.

However, the probability of detecting Q within some range — for example, that the atom decays between times 2 ms and 2.01 ms — is finite. We can therefore *discretize* the continuous variable: divide the range of values that Q can take into equal bins of width δQ. Then we define a discrete random variable $\tilde{Q}$ with possible values $\tilde{Q}_i$ equal to the central point of each bin, and the associated finite probability $\mathrm{pr}_{\tilde{Q}_i}$ that Q falls within that bin [Fig. B.4(a,b)]. As for any probability distribution, $\sum_i \mathrm{pr}_{\tilde{Q}_i} = 1$. Of course, the narrower the bin width we choose, the more precisely we describe the behavior of the continuous random variable.

The probability values associated with neighboring bins can be expected to be close to each other if the bin width is chosen sufficiently small. For atomic decay, for example, we can write $\mathrm{pr}_{[2.00\text{ ms},2.01\text{ ms}]} \approx \mathrm{pr}_{[2.01\text{ ms},2.02\text{ ms}]} \approx \frac{1}{2}\mathrm{pr}_{[2.00\text{ ms},2.02\text{ ms}]}$. In other words, for small bin widths, the quantity $\mathrm{pr}_{\tilde{Q}_i}/\delta Q$ is independent of δQ. Hence we can introduce the notion of the *probability density* or *continuous probability distribution*[1]:

$$\mathrm{pr}(Q) = \lim_{\delta Q \to 0} \frac{\mathrm{pr}_{\tilde{Q}_{i(Q)}}}{\delta Q}, \tag{B.10}$$

where $i(Q)$ is the number of the bin within which the value of Q is located and the limit is taken over a set of discretized probability distributions for Q. This probability density is the primary characteristic of a continuous random variable.

Note also that, because the discrete probability $\mathrm{pr}_{\tilde{Q}_{i(Q)}}$ is a dimensionless quantity, the dimension of a continuous probability density $\mathrm{pr}(Q)$ is always the reciprocal dimension of the corresponding random variable Q.

[1] Throughout this book, I use subscripts for discrete probabilities, such as in pr_i or $\mathrm{pr}_{\tilde{Q}_i}$, and parentheses for continuous probability densities, e.g., $\mathrm{pr}(Q)$.

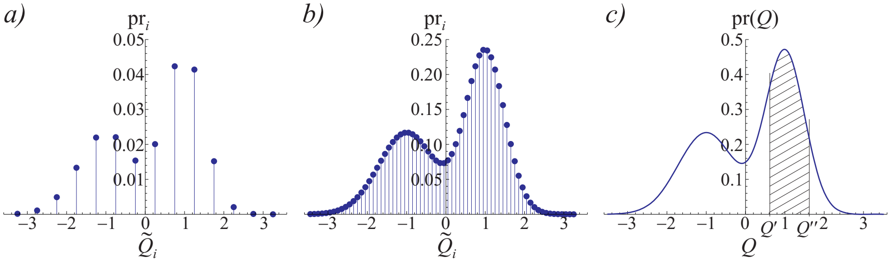

Fig. B.4 Continuous probability distribution. a), b) Discretization of the continuous random variable with bin widths $\delta Q = 0.5$ and 0.1, respectively. c) Continuous probability density. The probability of observing the variable in a range between Q' and Q'' is $\int_{Q'}^{Q''} \mathrm{pr}(Q)\mathrm{d}Q$. Note the variation of the vertical scales in the three plots.

Exercise B.16. For a continuous random variable with probability density $\mathrm{pr}(Q)$, show that:

a) the probability of observin the variable in the range between Q' and Q'' is

$$\mathrm{pr}_{[Q',Q'']} = \int_{Q'}^{Q''} \mathrm{pr}(Q)\mathrm{d}Q; \tag{B.11}$$

b) the probability density function is *normalized*:

$$\int_{-\infty}^{+\infty} \mathrm{pr}(Q)\mathrm{d}Q = 1; \tag{B.12}$$

c) the expectation value of Q is given by

$$\langle Q \rangle = \int_{-\infty}^{+\infty} Q\mathrm{pr}(Q)\mathrm{d}Q; \tag{B.13}$$

d)§ the variance of Q is given by

$$\left\langle \Delta Q^2 \right\rangle = \int_{-\infty}^{+\infty} (Q - \langle Q \rangle)^2 \mathrm{pr}(Q)\mathrm{d}Q = \left\langle Q^2 \right\rangle - \langle Q \rangle^2 . \tag{B.14}$$

Exercise B.17. Find the probability density, expectation and root mean square uncertainty for the decay time t of a radioactive nucleus with half-life $\tau = 1$ ms.

A probability density that frequently occurs in nature is the *Gaussian*, or *normal* distribution:

$$G_b(x) = \frac{1}{b\sqrt{\pi}} e^{-x^2/b^2}, \tag{B.15}$$

where b is the *width* of the Gaussian distribution (Fig. B.5). Typically, the Gaussian distribution governs physical quantities that are affected by multiple small random effects that add up[2]. Examples include:

- the position of a particle subjected to Brownian motion;
- the time shown by a clock affected by random fluctuations of the temperature in the room;
- the component of the velocity of a gas molecule along a particular axis.

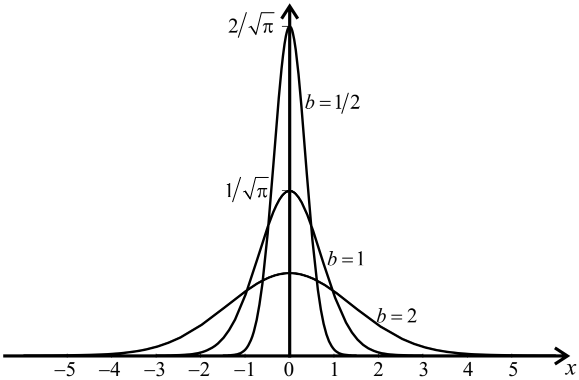

Fig. B.5 Normalized Gaussian functions of different widths.

Exercise B.18. For a Gaussian distribution $G_b(x-a)$, show the following:

a) Normalization holds, i.e.,

$$\int_{-\infty}^{+\infty} G_b(x)dx = 1. \tag{B.16}$$

Note that Eq. (B.17) also holds for complex b, as long as $\mathrm{Re}(b) > 0$.

b) The mean equals $\langle x \rangle = a$.

c) The variance is $\langle \Delta x^2 \rangle = b^2/2$.

Hint: use

[2] The rigorous formulation of this statement is called the *central limit theorem*.

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/b^2}\,dx = b\sqrt{\pi}; \tag{B.17}$$

$$\int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-x^2/b^2}\,dx = \frac{b^3\sqrt{\pi}}{2}. \tag{B.18}$$