1.3 PROBABILITY

Statistical Symbols Table

Statistics Symbol Sheet (greeks letters included)

PRIMERA PARTE: Distribuciones discretas

Because of the statistical interpretation, **probability** plays a central role in quantum mechanics, so I digress now for a brief discussion of the theory of probability. It is mainly a question of introducing some notation and terminology, and I shall do it in the context of a simple example.

Imagine a room containing 14 people, whose ages are as follows:

one person aged 14

one person aged 15

three people aged 16

two people aged 22

two people aged 24

five people aged 25.

If we let $N(j)$ represent the number of people of age j, then

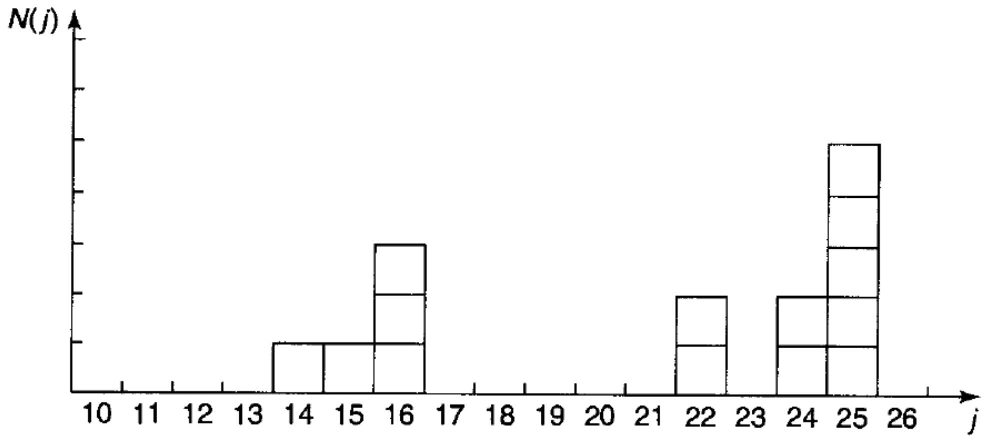


Figure 1.4: Histogram showing the number of people, $N(j)$, with age j, for the example in Section 1.3.

$$N(14) = 1$$
$$N(15) = 1$$
$$N(16) = 3$$
$$N(22) = 2$$
$$N(24) = 2$$
$$N(25) = 5$$

while $N(17)$, for instance, is zero. The *total* number of people in the room is

$$N = \sum_{j=0}^{\infty} N(j). \qquad [1.4]$$

(In this instance, of course, $N = 14$.) Figure 1.4 is a histogram of the data. The following are some questions one might ask about this distribution.

Question 1. If you selected one individual at random from this group, what is the **probability** that this person's age would be 15? *Answer:* One chance in 14, since there are 14 possible choices, all equally likely, of whom only one has this particular age. If $P(j)$ is the probability of getting age j, then $P(14) = 1/14$, $P(15) = 1/14$, $P(16) = 3/14$, and so on. In general,

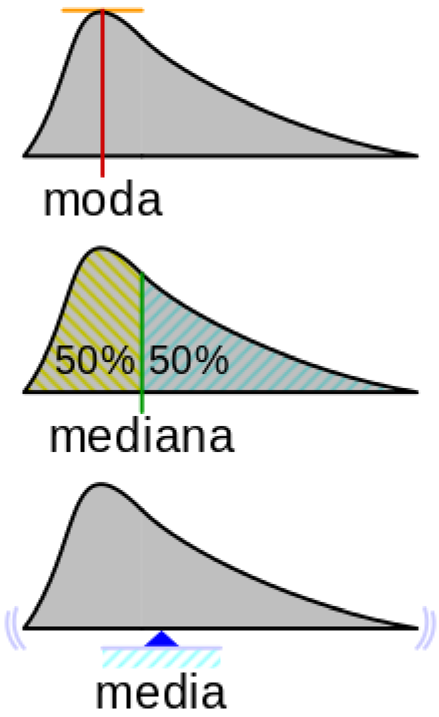
$$P(j) = \frac{N(j)}{N}. \qquad [1.5]$$

Notice that the probability of getting *either* 14 *or* 15 is the *sum* of the individual probabilities (in this case, 1/7). In particular, the sum of *all* the probabilities is 1—you're *certain* to get *some* age:

$$\sum_{j=1}^{\infty} P(j) = 1. \qquad [1.6]$$

Question 2. What is the **most probable** age? *Answer:* 25, obviously; five people share this age, whereas at most three have any other age. In general, the most probable j is the j for which $P(j)$ is a maximum.

Question 3. What is the **median** age? *Answer:* 23, for 7 people are younger than 23, and 7 are older. (In general, the median is that value of j such that the probability of getting a larger result is the same as the probability of getting a smaller result.)



Visualización geométrica de la moda, la mediana y de la media de una función arbitraria de densidad de probabilidad.

Question 4. What is the **average** (or **mean**) age? *Answer:*

$$\frac{(14)+(15)+3(16)+2(22)+2(24)+5(25)}{14}=\frac{294}{14}=21.$$

In general, the average value of j (which we shall write thus: $\langle j \rangle$) is given by

$$\langle j \rangle = \frac{\sum jN(j)}{N} = \sum_{j=0}^{\infty} jP(j). \qquad [1.7]$$

Notice that there need not be anyone with the average age or the median age—in this example nobody happens to be 21 or 23. In quantum mechanics the average is usually the quantity of interest; in that context it has come to be called the **expectation value**. It's a misleading term, since it suggests that this is the outcome you would be most likely to get if you made a single measurement (*that* would be the *most probable value*, not the average value)—but I'm afraid we're stuck with it.

Question 5. What is the average of the *squares* of the ages? *Answer:* You could get $14^2 = 196$, with probability 1/14, or $15^2 = 225$, with probability 1/14, or $16^2 = 256$, with probability 3/14, and so on. The average, then, is

$$\langle j^2 \rangle = \sum_{j=0}^{\infty} j^2 P(j). \qquad [1.8]$$

SEGUNDA PARTE: Distribuciones continuas

In general, the average value of some *function* of j is given by

$$\langle f(j)\rangle = \sum_{j=0}^{\infty} f(j)P(j). \qquad [1.9]$$

(Equations 1.6, 1.7, and 1.8 are, if you like, special cases of this formula.) *Beware:* The average of the squares ($\langle j^2\rangle$) is *not* ordinarily equal to the square of the average ($\langle j\rangle^2$). For instance, if the room contains just two babies, aged 1 and 3, then $\langle x^2\rangle = 5$, but $\langle x\rangle^2 = 4$.

$$< j > = \sum_{j=0}^{\infty} jP(j) = 1(0.5) + 3(0.5) = 2$$

$$< j >^2 = \left(\sum_{j=0}^{\infty} jP(j)\right)^2 = \left(1(0.5) + 3(0.5)\right)^2 = (2)^2 = 4$$

$$< j^2 > = \sum_{j=0}^{\infty} j^2P(j) = 1^2(0.5) + 3^2(0.5) = 1(0.5) + 9(0.5) = 0.5 + 4.5 = 5$$

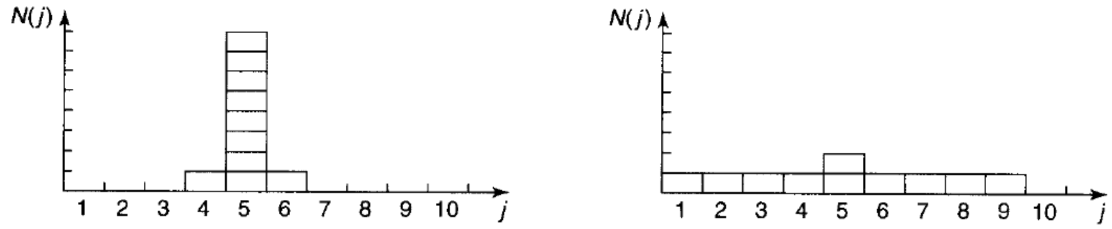


Figure 1.5: Two histograms with the same median, same average, and same most probable value, but different standard deviations.

Now, there is a conspicuous difference between the two histograms in Figure 1.5, even though they have the same median, the same average, the same most probable value, and the same number of elements: The first is sharply peaked about the average value, whereas the second is broad and flat. (The first might represent the age profile for students in a big-city classroom, and the second the pupils in a one-room schoolhouse.) We need a numerical measure of the amount of "spread" in a distribution, with respect to the average. The most obvious way to do this would be to find out how far each individual deviates from the average,

$$\Delta j = j - \langle j \rangle, \qquad [1.10]$$

and compute the average of Δj. Trouble is, of course, that you get *zero*, since, by the nature of the average, Δj is as often negative as positive:

$$< \Delta j > = \sum (j - < j >) P(j) = \sum \left(jP(j) - < j > P(j) \right) = \sum jP(j) - \sum < j > P(j)$$

$$< \Delta j > = \sum jP(j) - < j > \sum P(j)$$

De [1.7] se sabe que $\sum jP(j) = < j >$, entonces:

$$< \Delta j > = < j > - < j > \sum P(j)$$

De [1.6] se sabe que $\sum P(j) = 1$, entonces:

$$< \Delta j > = < j > - < j > (1) = < j > - < j > = 0$$

(Note that $\langle j \rangle$ is constant—it does not change as you go from one member of the sample to another—so it can be taken outside the summation.) To avoid this irritating problem, you might decide to average the *absolute value* of Δj. But absolute values are nasty to work with; instead, we get around the sign problem by *squaring* before averaging:

$$\sigma^2 \equiv \langle (\Delta j)^2 \rangle. \qquad [1.11]$$

This quantity is known as the **variance** of the distribution; σ itself (the square root of the average of the square of the deviation from the average—gulp!) is called the **standard deviation**. The latter is the customary measure of the spread about $\langle j \rangle$.

There is a useful little theorem involving standard deviations:

$$\sigma^2 = \langle (\Delta j)^2 \rangle = \sum (\Delta j)^2 P(j) = \sum (j - \langle j \rangle)^2 P(j)$$

$$= \sum (j^2 - 2j\langle j \rangle + \langle j \rangle^2) P(j)$$

$$= \sum j^2 P(j) - 2\langle j \rangle \sum jP(j) + \langle j \rangle^2 \sum P(j)$$

$$= \langle j^2 \rangle - 2\langle j \rangle \langle j \rangle + \langle j \rangle^2,$$

or

$$\sigma^2 = \langle j^2 \rangle - \langle j \rangle^2. \qquad [1.12]$$

Equation 1.12 provides a faster method for computing σ: Simply calculate $\langle j^2 \rangle$ and $\langle j \rangle^2$, and subtract. Incidentally, I warned you a moment ago that $\langle j^2 \rangle$ is not, in general,

equal to $\langle j \rangle^2$. Since σ^2 is plainly nonnegative (from its definition in Equation 1.11), Equation 1.12 implies that

$$\sigma^2 = < (\Delta j)^2 > = < j^2 > - < j >^2 \geq 0$$

$$< j^2 > \geq < j >^2$$

$$\langle j^2 \rangle \geq \langle j \rangle^2, \qquad [1.13]$$

and the two are equal only when $\sigma = 0$, which is to say, for distributions with no spread at all (every member having the same value).

So far, I have assumed that we are dealing with a *discrete* variable—that is, one that can take on only certain isolated values (in the example, j had to be an integer, since I gave ages only in years). But it is simple enough to generalize to *continuous* distributions. If I select a random person off the street, the probability that her age is *precisely* 16 years, 4 hours, 27 minutes, and 3.3333 seconds is *zero*. The only sensible thing to speak about is the probability that her age lies in some *interval*—say, between 16 years, and 16 years plus one day. If the interval is sufficiently short, this probability is proportional to the length of the interval. For example, the chance that her age is between 16 and 16 plus *two* days is presumably twice the probability that it is between 16 and 16 plus *one* day. (Unless, I suppose, there was some extraordinary baby boom 16 years ago, on exactly those days—in which case we have chosen an interval too long for the rule to apply. If the baby boom lasted six hours, we'll take intervals of a second or less, to be on the safe side. Technically, we're talking about infinitesimal intervals.) Thus

$$\left\{ \begin{array}{l} \text{probability that individual (chosen at random)} \\ \text{lies between } x \text{ and } (x + dx) \end{array} \right\} = \rho(x)\, dx. \qquad [1.14]$$

The proportionality factor, $\rho(x)$, is often loosely called "the probability of getting x," but this is sloppy language; a better term is **probability density**. The probability that x lies between a and b (a *finite* interval) is given by the integral of $\rho(x)$:

$$P_{ab} = \int_a^b \rho(x)\,dx, \qquad [1.15]$$

and the rules we deduced for discrete distributions translate in the obvious way:

$$\int_{-\infty}^{+\infty} \rho(x)\,dx = 1, \qquad [1.16]$$

$$\langle x \rangle = \int_{-\infty}^{+\infty} x\rho(x)\,dx, \qquad [1.17]$$

$$\langle f(x) \rangle = \int_{-\infty}^{+\infty} f(x)\rho(x)\,dx, \qquad [1.18]$$

$$\sigma^2 \equiv \langle (\Delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2. \qquad [1.19]$$

---

[7]The role of measurement in quantum mechanics is so critical and so bizarre that you may well be wondering what precisely *constitutes* a measurement. Does it have to do with the interaction between a microscopic (quantum) system and a macroscopic (classical) measuring apparatus (as Bohr insisted), or is it characterized by the leaving of a permanent "record" (as Heisenberg claimed), or does it involve the intervention of a conscious "observer" (as Wigner proposed)? I'll return to this thorny issue in the Afterword; for the moment let's take the naive view: A measurement is the kind of thing that a scientist does in the laboratory, with rulers, stopwatches, Geiger counters, and so on.