1 BOB makes his living writing answers for the back of mathematics textbooks. Unfortunately, BOB is not terribly careful, so each of his answers is correct only 90% of the time. One particular section has 20 problems that BOB must solve for the answer section. Find the probability that BOB gets at least 18 of these answers correct.

Solution: This is a Bernoulli trial: each problem is an independent even with probability of "success" (BOB getting the answer correct) $p = 0.90$. Since there are $n = 20$ problems, the probability that BOB gets k of them correct is

$$P(X = k \text{ “successes”}) = b(20, k; 0.90) = \binom{20}{k}(0.90)^k(0.10)^{20-k}.$$

Thus the probability that BOB gets at least 18 answers correct is

$$\begin{aligned} P(X \geq 18) &= P(X = 18) + P(X = 19) + P(X = 20) \\ &= b(20, 18; 0.90) + b(20, 19; 0.90) + b(20, 20; 0.90) \\ &= \binom{20}{18}(0.90)^{18}(0.10)^2 + \binom{20}{19}(0.90)^{19}(0.10) + \binom{20}{20}(0.90)^{20} \\ &\approx 0.6769268. \end{aligned}$$

Thus BOB gets at least 18 out of 20 answers correct about 67.7% of the time (slightly more than two-thirds of the time).

2 It is well known that the measure of Intelligence Quotient (IQ) has a normal distribution among the general population. The mean IQ for the general population is $\mu = 100$ and the standard deviation is $\sigma = 16$.

(a) What percentage of the population has an IQ score between 84 and 132?

Solution: Let X be the random variable describing the IQ of the general population. We are told that X has a random distribution with mean $\mu = 100$ and standard deviation $\sigma = 16$, and we are asked for $P(84 \leq X \leq 132)$. We calculate this in the usual way – transforming the random variable X into the standard normal distribution Z:

$$\begin{aligned} P(84 \leq X \leq 132) &= P\left(\frac{84 - 100}{16} \leq Z \leq \frac{132 - 100}{16}\right) \\ &= P\left(-1 \leq Z \leq 2\right). \end{aligned}$$

Now the trick is to write this in terms of Z-values; that is, in terms of probabilities $P(0 \leq Z \leq z)$ that we can look up on tables. We do this using the symmetry of the normal curve:

$$\begin{aligned} P(84 \leq X \leq 132) &= P\left(-1 \leq Z \leq 2\right) \\ &= P(-1 \leq Z \leq 0) + P(0 \leq Z \leq 2) \\ &= P(0 \leq Z \leq 1) + P(0 \leq Z \leq 2) \qquad \text{by symmetry} \\ &= 0.3413 + 0.4772 = 0.8185 \qquad \text{via the table.} \end{aligned}$$

Thus about 81.85% of the general population has an IQ in the range 84 to 132.

(b) Mensa is an international society whose membership is limited to persons having IQs above the general population's 98th percentile (that is, people with IQ scores in the top two percent). What is the lowest IQ score that will qualify a person to belong to Mensa?

Solution: We're asked for to find x for which $P(X \geq x) = 0.02$ (or, equivalently, $P(X \leq x) = 0.98$). Here X is a normal random variable with mean $\mu = 100$ and standard deviation $\sigma = 16$. Thus we can change this into the standard normal distribution Z in the usual way:

$$0.02 = P(X \geq x) = P\left(Z \geq \frac{x-\mu}{\sigma}\right) = P\left(Z \geq \frac{x-100}{16}\right) = P(Z \geq z),$$

where $z = \frac{x-100}{16}$. Thus we want to find z where

$$P(0 \leq Z \leq z) = P(Z \geq 0) - P(Z \geq z) = 0.5 - 0.02 = 0.4800.$$

To find this, we look at our table. Alas, nothing on the table gives us a Z-value of precisely 0.4800. We can interpolate, however, to find such a z:

z	2.05	z	2.06
Z-value	0.4798	0.4800	0.4803

Often what's done in this situation is we interpolate between 2.05 and 2.06 to find the z with Z-value 0.4800:

$$\frac{\Delta z}{\Delta \text{ Value}} = \frac{z - 2.05}{0.4800 - 0.4798} = \frac{2.06 - 2.05}{0.4803 - 0.4798} \quad \text{or} \quad \frac{z-2.05}{0.0002} = \frac{0.01}{0.0005}.$$

Thus $z = 2.05 + \frac{0.01}{0.0005}(0.0002) = 2.054$. Solving for x in $z = \frac{x-100}{16}$, we get the following possible values:

z	x
2.05	132.8
2.054	132.864
2.06	132.96

So roughly speaking the answer is always around an IQ of 133.

3 Dr. Patty O'Furniture's patients must really like her. From the time they arrive at her office, they expect to wait 90 minutes before being taken in for their appointment! Assume that patient wait time (measured from arrival in the office) is an exponentially distributed random variable. Dr. O'Furniture would like to post a sign in her waiting room informing her patients of the likelihood of various waiting times. To assist her, please find the following probabilities:

(a) The probability that a patient will wait more than three hours.

(b) The probability that a patient will wait less than 45 minutes.

(c) The probability that a patient will wait between 45 and 180 minutes.

You may find these probabilities in any order you wish, but please label your solutions clearly.

Solution: A probability using an exponential distribution is of the form

$$P(a \leq X \leq b) = \int_a^b \lambda e^{-\lambda x}\, dx,$$

where λ is determined by the mean. For an exponential random variable, the mean is $E(X) = \frac{1}{\lambda}$. Since we are told that $E(X) = 90$ minutes, we get that $\lambda = \frac{1}{90}$. Thus the probabilities are

$$P(a \leq X \leq b) = \int_a^b \frac{1}{90} e^{-\frac{1}{90}x}\, dx = -e^{-\frac{1}{90}x}\Big|_a^b.$$

We'll use this computation three times:

(a) This question asks for $P(X \geq 180)$ (we've used minutes, remember, not hours). This can be written as $P(180 \leq X \leq \infty)$ (so we get an explicit value for b). This is an improper integral, and so we must use the usual limit:

$$\begin{aligned}
P(X \geq 180) &= \int_{180}^{\infty} \frac{1}{90} e^{-\frac{1}{90}x}\, dx = \lim_{b\to\infty} \int_{180}^{b} \frac{1}{90} e^{-\frac{1}{90}x}\, dx \\
&= \lim_{b\to\infty} \left(-e^{-\frac{1}{90}x}\Big|_{180}^{b} \right) \\
&= \lim_{b\to\infty} \left[\left(-e^{-\frac{1}{90}b}\right) - \left(-e^{-\frac{1}{90}(180)}\right) \right] \\
&= \lim_{b\to\infty} \left(-e^{-\frac{1}{90}b} + e^{-2} \right) \\
&= 0 + e^{-2} \approx 0.1353 \quad \text{or } 13.53\%.
\end{aligned}$$

Thus $P(X \geq 180) = e^{-2} \approx 13.53\%$.

(b) This asks for $P(0 \leq X \leq 45)$. This is a straightforward computation:

$$\begin{aligned}
P(0 \leq X \leq 45) &= \int_0^{45} \frac{1}{90} e^{-\frac{1}{90}x}\, dx \\
&= \left(-e^{-\frac{1}{90}x}\Big|_0^{45} \right) \\
&= \left(-e^{-\frac{1}{90}(45)}\right) - \left(-e^{-\frac{1}{90}(0)}\right) \\
&= 1 - e^{-1/2} \approx 0.3935 \quad \text{or } 39.35\%.
\end{aligned}$$

Thus $P(0 \leq X \leq 45) = 1 - e^{-1/2} \approx 39.35\%$.

(c) This asks for $P(45 \le X \le 180)$. This is again a straightforward computation, very similar to the previous parts:

$$\begin{aligned} P(45 \le X \le 180) &= \int_{45}^{180} \frac{1}{90} e^{-\frac{1}{90}x}\, dx \\ &= \left(-e^{-\frac{1}{90}x} \Big|_{45}^{180} \right) \\ &= \left(-e^{-\frac{1}{90}(180)} \right) - \left(-e^{-\frac{1}{90}(45)} \right) \\ &= e^{-1/2} - e^{-2} \approx 0.4712 \quad \text{or } 47.12\%. \end{aligned}$$

Thus $P(45 \le X \le 180) = e^{-1/2} - e^{-2} \approx 47.12\%$.

Notice that

$$P(X \ge 180) + P(0 \le X \le 45) + P(45 \le X \le 180) = 1$$

which shows up as

$$e^{-2} + \left(1 - e^{-1/2}\right) + \left(e^{-1/2} - e^{-2}\right) = 1 \qquad \text{or} \qquad 13.53\% + 39.35\% + 47.12\% = 100\%.$$

Thus we could have computed one of (a), (b), or (c) from the other two.

4 Suppose $f(x) = \dfrac{k}{x^3}$ for $x \ge 1$.

(a) Find k so that $f(x)$ is a probability density function over the interval $[1, \infty)$ (or, if you prefer, $x \ge 1$).

Solution: To be a probability density function on $[1, \infty)$, $f(x)$ must satisfy two conditions: (1) $f(x) \ge 0$ for all $x \ge 1$, and (2) the total probability must be 1:

$$\int_1^\infty f(x)\, dx = 1.$$

The first condition simply says that $k \ge 0$. The second will determine k for us. We compute:

$$\begin{aligned} 1 &= \int_1^\infty f(x)\, dx = \int_1^\infty \frac{k}{x^3}\, dx \\ &= k \int_1^\infty x^{-3}\, dx \\ &= k \lim_{b\to\infty} \int_1^b x^{-3}\, dx \\ &= k \lim_{b\to\infty} \left(\frac{x^{-2}}{-2} \Big|_1^b \right) = k \lim_{b\to\infty} \left(-\frac{1}{2x^2} \Big|_1^b \right) \\ &= k \lim_{b\to\infty} \left[-\frac{1}{2b^2} - \left(-\frac{1}{2(1)^2} \right) \right] = k \lim_{b\to\infty} \left(-\frac{1}{2b^2} + \frac{1}{2} \right) \\ &= k \left(0 + \frac{1}{2} \right) = \frac{1}{2}k. \end{aligned}$$

Thus $\frac{1}{2}k = 1$, so $k = 2$.

(b) Suppose X is a random variable with probability density function $f(x)$ (as given in part (a)). Find $E(X)$, the expected value of X.

Solution: Recall that the expected value of X is given by the formula

$$E(X) = \int_a^b x f(x)\, dx.$$

Thus we simply must compute the integral

$$E(X) = \int_1^\infty x \frac{2}{x^3}\, dx.$$

Let's do this computation:

$$\begin{aligned} E(X) &= \int_1^\infty x \frac{2}{x^3}\, dx = \int_1^\infty \frac{2}{x^2}\, dx = 2\int_1^\infty x^{-2}\, dx \\ &= 2 \lim_{b\to\infty} \int_1^b x^{-2}\, dx \\ &= 2 \lim_{b\to\infty} \left(\frac{x^{-1}}{-1}\Big|_1^b \right) = 2 \lim_{b\to\infty} \left(-\frac{1}{x}\Big|_1^b \right) \\ &= 2 \lim_{b\to\infty} \left[-\frac{1}{b} - \left(-\frac{1}{1} \right) \right] = 2 \lim_{b\to\infty} \left(-\frac{1}{b} + 1 \right) \\ &= 2\,(-0+1) = 2. \end{aligned}$$

Thus the expected value is $E(X) = 2$.

5 (a) Find the *mean*, *median*, *variance* and *standard deviation* of the following data:

$$S_1 = \{\ -3,\ 1,\ 3,\ 9,\ 10\ \}.$$

Show enough of the details of your computations so that another person can follow them.

Solution: The mean of this data is simply

$$\mu = \frac{-3+1+3+9+10}{5} = \frac{20}{5} = 4.$$

The median is the middle value (since there are an odd number of data points – if there had been an even number, we would have had to average the two middle values). In this case the median is 3. The (population) variance we may compute as

$$\begin{aligned} \sigma^2 &= \frac{(-3-4)^2 + (1-4)^2 + (3-4)^2 + (9-4)^2 + (10-4)^2}{5} \\ &= \frac{(-7)^2 + (-3)^2 + (-1)^2 + (5)^2 + (6)^2}{5} \\ &= \frac{49+9+1+25+36}{5} \\ &= \frac{120}{5} = 24. \end{aligned}$$

From this we can find the standard deviation $\sigma = \sqrt{24} = 2\sqrt{6} \approx 4.8990$.

(b) Suppose the data set

$$S_2 = \{-1,\ 3,\ x,\ y,\ z\ \}$$

has mean zero, median zero, and variance $\sigma^2 = 6$. Find x, y, and z.

Solution: Since there are five data points, the median must be part of the data (and not the average of two data points). Let us call the median $x = 0$. The mean is zero; this means that

$$0 = \frac{-1+3+0+y+z}{5} \qquad \text{or} \qquad 0 = y+z+2 \qquad \text{or} \qquad y = -z-2.$$

Since the variance is $\sigma^2 = 6$, we get

$$6 = \frac{(-1-0)^2 + (3-0)^2 + (0-0)^2 + (y-0)^2 + (z-0)^2}{5}$$

or $30 = 1 + 9 + y^2 + z^2$ or $20 = y^2 + z^2$. Plugging in $y = -z - 2$ into this last equation we get

$$20 = (-z-2)^2 + z^2 \qquad \text{or} \qquad 2z^2 + 4z + 4 = 20 \qquad \text{or} \qquad z^2 + 2z - 8 = 0.$$

Factoring, we get $(z-2)(z+4) = 0$, so $z = 2$ or $z = -4$. When $z = 2$, $y = -z - 2 = -4$ and (similarly) when $z = -4$, $y = -z - 2 = 2$. Thus $\{x,\ y,\ z\} = \{0,\ 2,\ -4\}$ (in some order).