Chapter 5 Important continuous distributions
Bibliografie: Riley et al. (2006), Montgomery (2011)

Having discussed the most commonly encountered discrete probability distributions, we now consider some of the more important continuous probability distributions. These are summarised in table 1.2.

Table 1.2

Distribution	Probability law $f(x)$	MGF	$E[X]$	$V[X]$
Gaussian	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$e^{\mu t+\frac{1}{2}\sigma^2 t^2}$	μ	σ^2
Exponential	$\lambda e^{-\lambda x}$	$\frac{\lambda}{\lambda-t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma	$\frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}$	$\left(\frac{\lambda}{\lambda-t}\right)^r$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$
Chi-squared	$\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{1}{2}n\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}$	$\left(\frac{1}{1-2t}\right)^{n/2}$	n	$2n$
Uniform	$\frac{1}{b-a}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$

5.1 The Gaussian distribution

By far the most important continuous probability distribution is the *Gaussian* or *normal* distribution. Many random variables of interest, in all areas of the physical sciences and beyond, are described either exactly or approximately by a Gaussian distribution. Moreover, the Gaussian distribution can be used to approximate other, more complicated, probability distributions.

The *probability density function for a Gaussian distribution* of a random variable *X*, with mean $E[X]=\mu$ and $V[X]=\sigma^2$, is

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (5.1)$$

The factor $1/\sqrt{2\pi}$ arises from the normalisation of the distribution,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \tag{5.2}$$

The evaluation of this type of integral is $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ (prin ridicare la patrat si integrare in coordonate polare). The Gaussian distribution is *symmetric* about the point $x = \mu$ and has the characteristic 'bell' shape shown in figure 5.1.

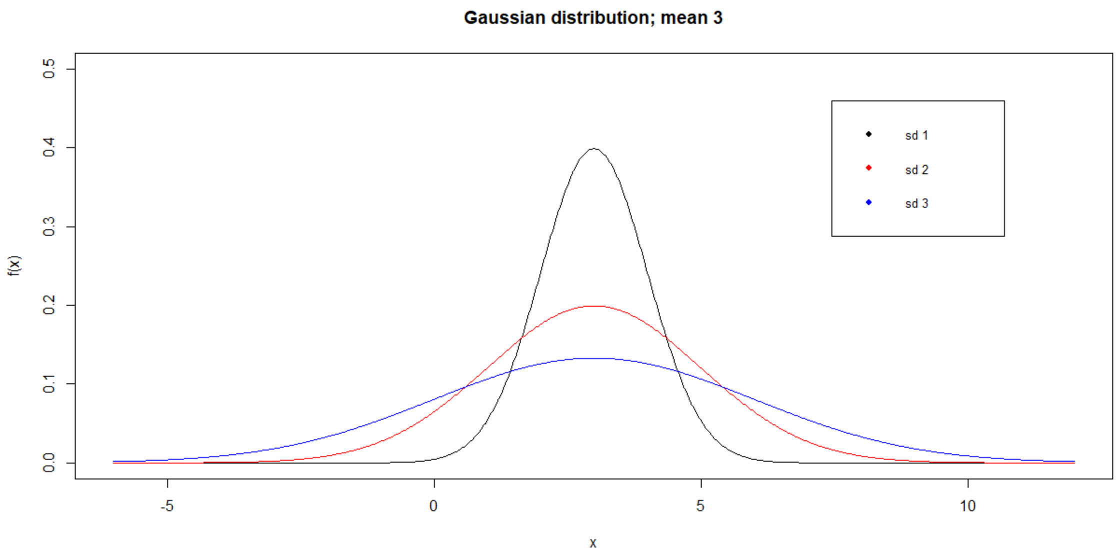

Figure 5.1 The Gaussian or normal distribution for mean $\mu = 3$ and various values of the standard deviation σ

The *width* of the curve is described by the *standard deviation* σ: if σ is large then the curve is broad, and if σ is small then the curve is narrow (see the figure). At $x = \mu \pm \sigma$, $f(x)$ falls to $e^{-1/2} \approx 0.61$ of its peak value; these points are points of inflection, where $d^2 f / dx^2 = 0$.

Indeed,

$$f'(x) = -\frac{1}{\sigma^2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \frac{x-\mu}{\sigma} \tag{5.3}$$

$$f''(x) = -\frac{1}{\sigma^2\sqrt{2\pi}} \left(\frac{1}{\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} - \frac{1}{2} 2 \left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right)$$

$$f''(x) = -\frac{1}{\sigma^3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \left(1 - \left(\frac{x-\mu}{\sigma}\right)^2\right) \tag{5.4}$$

$$d^2 f / dx^2 = 0 \quad \Rightarrow \quad 1 - \left(\frac{x-\mu}{\sigma}\right)^2 = 0 \quad \Rightarrow \quad x = \mu \pm \sigma$$

$$f(\mu \pm \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu\pm\sigma-\mu}{\sigma}\right)^2} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}}$$

When a random variable *X* follows a *Gaussian distribution* with mean μ and variance σ^2, we write $X \sim N(\mu, \sigma^2)$.

The *effects of changing* μ and σ are only to *shift* the curve along the *x*-axis or to *broaden* or *narrow* it, respectively. Thus all Gaussians are equivalent in that a change of origin and scale can reduce them to a *standard form*. We consider the random variable $Z = \frac{X-\mu}{\sigma}$, for which the PDF takes the form

$$x = \sigma z + \mu \qquad g(z) = f(x(z))\left|\frac{dx}{dz}\right| = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\sigma z+\mu-\mu}{\sigma}\right)^2} |\sigma| = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tag{5.5}$$

which is called the *standard Gaussian distribution* and has mean $\mu = 0$ and variance $\sigma^2 = 1$. The random variable *Z* is called the *standard variable*.

The *cumulative probability function* for a Gaussian distribution is

$$F(x) = \Pr(X < x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du \tag{5.6}$$

where *u* is a (dummy) integration variable. This integral cannot be evaluated analytically. It is a standard practice to tabulate values of the cumulative probability function for the standard Gaussian distribution (see figure 5.2), i.e.

$$\Phi(z) = \Pr(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}} du \qquad (5.7)$$

It is usual only to tabulate $\Phi(z)$ for $z > 0$, since it can be seen easily, from figure 5.2 and the symmetry of the Gaussian distribution, that $\Phi(-z) = 1 - \Phi(z)$; see table 5.1. Using such a table it is then straightforward to evaluate the probability that *Z* lies in a given range of *z*-values. For example, for *a* and *b* constant,

$$\Pr(Z < a) = \Phi(a) \qquad (5.8)$$

$$\Pr(Z > a) = 1 - \Phi(a) \qquad (5.9)$$

$$\Pr(a < Z \le b) = \Phi(b) - \Phi(a) \qquad (5.10)$$

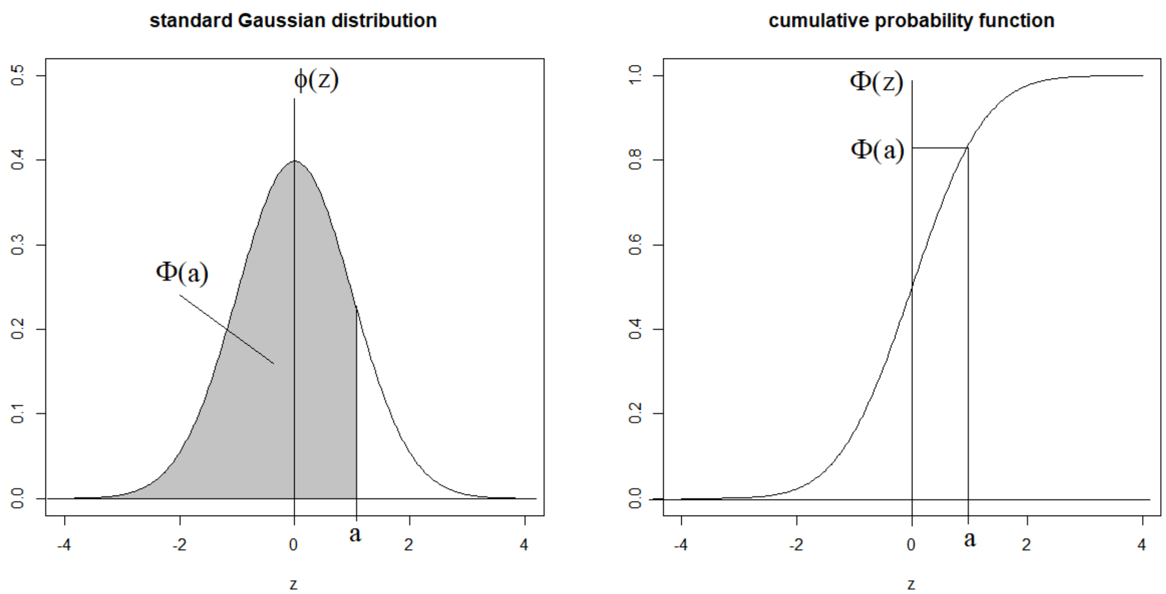

Figure 5.2 On the left, the standard Gaussian distribution $\phi(z)$; the shaded area gives $\Pr(Z < a) = \Phi(a)$. On the right, the cumulative probability function $\Phi(z)$ for a standard Gaussian distribution $\phi(z)$.

Remembering that $Z = \frac{X - \mu}{\sigma}$ and

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{5.11}$$

Indeed, $\Phi\left(\frac{x-\mu}{\sigma}\right) = \Pr\left(Z < \frac{x-\mu}{\sigma}\right) = \Pr(Z\sigma + \mu < x) = \Pr(X < x) = F(x)$.

So, we may also calculate the probability that the original random variable X lies in a given x-range. For example,

$$\Pr(a < X \le b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \tag{5.12}$$

$$= F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \tag{5.13}$$

Table 5.1 Cumulative Standard Normal Distribution

Φ(z)	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
	...									

Exercise 1. If X is described by a Gaussian distribution of mean μ and variance σ^2, calculate the probabilities that X lies within 1σ, 2σ and 3σ of the mean.

From (5.13)

$$\Pr(\mu - n\sigma < X \le \mu + n\sigma) = \Phi(n) - \Phi(-n) = \Phi(n) - [1 - \Phi(n)] = 2\Phi(n) - 1$$

and so from table 5.1, or with R

$$\Pr(\mu - \sigma < X \le \mu + \sigma) = 2\Phi(1) - 1 = 0.6826 \approx 68.3\%$$

$$\Pr(\mu - 2\sigma < X \le \mu + 2\sigma) = 2\Phi(2) - 1 = 0.9544 \approx 95.4\%$$

$$\Pr(\mu - 3\sigma < X \le \mu + 3\sigma) = 2\Phi(3) - 1 = 0.9974 \approx 99.7\%$$

Thus we expect X to be distributed in such a way that about two thirds of the values will lie between $\mu - \sigma$ and $\mu + \sigma$, 95% will lie within 2σ of the mean and 99.7% will lie within 3σ of the mean. These limits are called the one-, two- and three-sigma limits respectively; it is important to note that they are independent of the actual values of the mean and variance.

Exercise 2. Sawmill (fabrica de cherestea) *A* produces boards (scanduri) whose lengths are Gaussian distributed with mean 209.4 cm and standard deviation 5.0 cm. A board is accepted if it is longer than 200 cm but is rejected otherwise. Show that 3% of boards are rejected.

Let *X* = length of boards from *A*, so that $X \sim N\left(209.4,(5.0)^2\right)$ and

$$\Pr\left(X<200\right)=\Phi\left(\frac{200-\mu}{\sigma}\right)=\Phi\left(\frac{200-209.4}{5.0}\right)=\Phi\left(-1.88\right)$$

$$\Pr\left(X<200\right)=1-\Phi\left(1.88\right)=1-0.9699=0.0301$$

i.e. 3.0% of boards are rejected.

Exercise 3. The time taken for a computer ‘packet’ to travel from Cambridge UK to Cambridge MA is Gaussian distributed. 6.8% of the packets take over 200 ms to make the journey, and 3.0% take under 140 ms. Find the mean and standard deviation of the distribution.

Let *X* = journey time in ms; we are told that $X \sim N\left(\mu,\sigma^2\right)$ where μ and σ are unknown. Since 6.8% of journey times are longer than 200 ms,

$$\Pr\left(X>200\right)=1-\Pr\left(X<200\right)=1-\Phi\left(\frac{200-\mu}{\sigma}\right)=0.068$$

$$\Phi\left(\frac{200-\mu}{\sigma}\right)=1-0.068=0.932$$

Using table 5.1, we have therefore

$$\frac{200-\mu}{\sigma}=1.49 \qquad (5.14)$$

Also, 3.0% of journey times are under 140 ms, so

$$\Pr\left(X<140\right)=\Phi\left(\frac{140-\mu}{\sigma}\right)=0.030$$

Now using $\Phi(-z)=1-\Phi(z)$ gives

$$\Phi\left(\frac{\mu-140}{\sigma}\right)=1-0.030=0.970$$

Using table 5.1, we find

$$\frac{\mu-140}{\sigma}=1.88 \tag{5.15}$$

Solving the simultaneous equations (5.14) and (5.15) gives $\mu=173.5$, $\sigma=17.8$.

The moment generating function for the Gaussian distribution

$$M_X(t)=E\left[e^{tX}\right]=\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[tx-\frac{(x-\mu)^2}{2\sigma^2}\right]dx \tag{5.16}$$

Since

$$tx-\frac{(x-\mu)^2}{2\sigma^2}=\frac{2tx\sigma^2-x^2+2x\mu-\mu^2}{2\sigma^2}$$

$$=\frac{2x\left(t\sigma^2+\mu\right)-x^2-\left(t\sigma^2+\mu\right)^2+\left(t\sigma^2+\mu\right)^2-\mu^2}{2\sigma^2}$$

$$=\frac{-\left[x-\left(t\sigma^2+\mu\right)\right]^2+2t\sigma^2\mu+t^2\sigma^4}{2\sigma^2}$$

$$M_X(t)=E\left[e^{tX}\right]=\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[t\mu+\frac{\sigma^2t^2}{2}-\frac{\left[x-\left(\mu+\sigma^2t\right)\right]^2}{2\sigma^2}\right]dx$$

$$M_X(t)=c\exp\left(\mu t+\frac{\sigma^2t^2}{2}\right)$$

where the final equality is established by completing the square in the argument of the exponential and writing

$$c=\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{\left[x-\left(\mu+\sigma^2t\right)\right]^2}{2\sigma^2}\right]dx$$

The final integral is simply the normalization integral for the Gaussian distribution, and so $c = 1$ and the MGF is given by

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \qquad (5.17)$$

$$M'_X(t) = \left(\mu + \sigma^2 t\right) e^{\mu t + \frac{1}{2}\sigma^2 t^2} \qquad E[X] = M'_X(0) = \mu$$

$$M''_X(t) = \left[\sigma^2 + \left(\mu + \sigma^2 t\right)^2\right] e^{\mu t + \frac{1}{2}\sigma^2 t^2} \qquad E\left[X^2\right] = M''_X(0) = \sigma^2 + \mu^2$$

$$V[X] = M''_X(0) - \left[M'_X(0)\right]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

Gaussian approximation to the binomial distribution

We may consider the Gaussian distribution as the limit of the binomial distribution when the number of trials $n \to \infty$ but the probability of a success p remains finite, so that $np \to \infty$ also. (This contrasts with the Poisson distribution, which corresponds to the limit $n \to \infty$ and $p \to 0$ with $np = \lambda$ remaining finite.) In other words, a Gaussian distribution results when an experiment with a finite probability of success is repeated a large number of times. To see how this Gaussian limit arises look Riley et al. (2006). The binomial probability function gives the probability of x successes in n trials as

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \qquad (5.18)$$

Taking the limit as $n \to \infty$ (and $x \to \infty$) we may approximate (Riley et al. (2006))

$$f(x) \approx \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{np(1-p)}} \exp\left[-\frac{1}{2}\frac{(x-np)^2}{np(1-p)}\right], \qquad (5.19)$$

which is of Gaussian form with $\mu = np$ and $\sigma = \sqrt{np(1-p)}$.

Thus we see that the value of the Gaussian probability density function $f(x)$ is a good approximation to the probability of obtaining x successes in n trials. This approximation is actually very good even for relatively small n. For example, if $n = 10$ and $p = 0.6$ then the Gaussian approximation to the binomial distribution have $\mu = 10 \times 0.6 = 6$ and $\sigma = \sqrt{10 \times 0.6(1-0.6)} = 1.549$. The probability functions $f(x)$ for the binomial and associated Gaussian distributions for these parameters are given in table 5.2, and it can be seen that the Gaussian approximation is a good one.

Table 5.2 Comparison of the binomial distribution for $n=10$ and $p=0.6$ with its Gaussian approximation.

x	$f(x)$ binomial	$f(x)$ Gaussian
0	0.0001	0.0001
1	0.0016	0.0014
2	0.0106	0.0092
3	0.0425	0.0395
4	0.1115	0.1119
5	0.2007	0.2091
6	0.2508	0.2575
7	0.2150	0.2091
8	0.1209	0.1119
9	0.0403	0.0395
10	0.0060	0.0092

Since the *Gaussian distribution is continuous* and the *binomial distribution is discrete*, we should use the integral of $f(x)$ for the Gaussian distribution in the calculation of approximate binomial probabilities. We should apply a *continuity correction* so that the discrete integer *x* in the binomial distribution becomes the interval $[x-0.5, x+0.5]$ in the Gaussian distribution. Explicitly,

$$\Pr(X=x) \approx \frac{1}{\sigma\sqrt{2\pi}} \int_{x-0.5}^{x+0.5} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du \tag{5.20}$$

The Gaussian approximation is useful for estimating the binomial probability that *X* lies between the (integer) values x_1 and x_2,

$$\Pr(x_1 \le X \le x_2) \approx \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1-0.5}^{x_2+0.5} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du \tag{5.21}$$

Exercise 4. A manufacturer makes computer chips of which 10% are defective. For a random sample of 200 chips, find the approximate probability that more than 15 are defective.

We first define the random variable

X = number of defective chips in the sample,

which has a binomial distribution, $X \sim Bin(200, 0.1)$. Therefore, the mean and variance of this distribution are

$$E[X] = 200 \times 0.1 = 20 \quad \text{and} \quad V[X] = 200 \times 0.1(1-0.1) = 18$$

and we may approximate the binomial distribution with a Gaussian distribution such that $X \sim N(20,18)$. The standard variable is

$$Z = \frac{X-20}{\sqrt{18}}$$

and so, using $x = 15.5$ to allow for the continuity correction,

$$\Pr(X > 15.5) = \Pr\left(Z > \frac{15.5-20}{\sqrt{18}}\right) = \Pr(Z > -1.06) = \Pr(Z < 1.06) = 0.86$$

Gaussian approximation to the Poisson distribution

We first met the Poisson distribution as the limit of the binomial distribution for $n \to \infty$ and $p \to 0$, taken in such a way that $np = \lambda$ remains finite. Further, in the previous subsection, we considered the Gaussian distribution as the limit of the binomial distribution when $n \to \infty$ but p remains finite, so that $np \to \infty$ also. It should come as no surprise, that the Gaussian distribution can also be used to approximate the Poisson distribution when the mean λ becomes large. The probability function for the Poisson distribution is

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!} \tag{5.22}$$

It can be shown that (Riley et al. (2006))

$$f(x) \approx \frac{1}{\sqrt{2\pi\lambda}} e^{-\frac{(x-\lambda)^2}{2\lambda}} \tag{5.23}$$

which is the Gaussian distribution with $\mu = \lambda$ and $\sigma^2 = \lambda$.

The larger the value of λ, the better is the Gaussian approximation to the Poisson distribution; the approximation is reasonable even for $\lambda = 5$, but $\lambda \geq 10$ is safer. As in the case of the Gaussian approximation to the binomial distribution, a *continuity correction* is necessary since the Poisson distribution is discrete.

Exercise 5. E-mail messages are received by an author at an average rate of one per hour. Find the probability that in a day the author receives 24 messages or more.

We first define the random variable

X = number of messages received in a day.

Thus $E[X] = 1 \times 24 = 24$, and so $X \sim Po(24)$. Since $\lambda > 10$ we may approximate the Poisson distribution by $X \sim N(24, 24)$. Now the standard variable is

$$Z = \frac{X - 24}{\sqrt{24}} \ ,$$

and, using the continuity correction, we find

$$\Pr(X > 23.5) = \Pr\left(Z > \frac{23.5 - 24}{\sqrt{24}} \right) = \Pr(Z > -0.102) = \Pr(Z < 0.102) = 0.54$$

In fact, almost all probability distributions tend towards a Gaussian when the numbers involved become large – that this should happen is required by the *central limit theorem*, which we discuss in a next section.

Multiple Gaussian distributions

Suppose X and Y are independent Gaussian-distributed random variables, so that

$$X \sim N\left(\mu_1, \sigma_1^2\right) \quad \text{and} \quad Y \sim N\left(\mu_2, \sigma_2^2\right)$$

Let us consider the random variable $Z = X + Y$. The PDF for this random variable may be found using the MGFs. The MGFs of X and Y are

$$M_X(t) = e^{\mu_1 t + \frac{1}{2}\sigma_1^2 t^2} \quad , \quad M_Y(t) = e^{\mu_2 t + \frac{1}{2}\sigma_2^2 t^2}$$

Since X and Y are independent RVs, the MGF of $Z = X + Y$ is simply the product of $M_X(t)$ and $M_Y(t)$:

$$M_Z(t) = M_X(t) M_Y(t) = e^{\mu_1 t + \frac{1}{2}\sigma_1^2 t^2} e^{\mu_2 t + \frac{1}{2}\sigma_2^2 t^2}$$

$$= e^{(\mu_1 + \mu_2)t + \frac{1}{2}\left(\sigma_1^2 + \sigma_2^2\right)t^2} \quad ,$$

which we recognise as the MGF for a Gaussian with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$. Thus, Z is also Gaussian distributed: $Z \sim N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$.

Exercise 6. An executive travels home from her office every evening. Her journey consists of a train ride, followed by a bicycle ride. The time spent on the train is Gaussian distributed with mean 52 minutes and standard deviation 1.8 minutes, while the time for the bicycle journey is Gaussian distributed with mean 8 minutes and standard deviation 2.6 minutes. Assuming these two factors are independent, estimate the percentage of occasions on which the whole journey takes more than 65 minutes.

We first define the random variables

X = time spent on train, Y= time spent on bicycle,

so that, $X \sim N\left(52, (1.8)^2\right)$ and $Y \sim N\left(8, (2.6)^2\right)$. Since X and Y are independent, the total journey time $T = X + Y$ is distributed as

$$T \sim N\left(52 + 8, \quad (1.8)^2 + (2.6)^2\right) = N\left(60, \quad (3.16)^2\right)$$

The standard variable is thus

$$Z = \frac{T - 60}{3.16}$$

and the required probability is given by

$$\Pr(T > 65) = \Pr\left(Z > \frac{65 - 60}{3.16}\right) = \Pr(Z > 1.58) = 1 - \Pr(Z < 1.58) = 1 - 0.943 = 0.057$$

Thus the total journey time exceeds 65 minutes on 5.7% of occasions.

The above results may be extended. For example, if the random variables X_i, $i = 1, 2, \ldots, n$, are distributed as $X_i \sim N\left(\mu_i, \sigma_i^2\right)$ then the random variable $Z = \sum_i c_i X_i$ (where the c_i are constants) is distributed as $Z \sim N\left(\sum_i c_i \mu_i, \sum_i c_i^2 \sigma_i^2\right)$.

5.2 The log-normal distribution

If the random variable X follows a Gaussian distribution then the variable $Y = e^X$ is described by a *log-normal* distribution. If X can take values in the range $-\infty$ to ∞, then Y will lie between 0 and ∞. The probability density function for Y is

$$g(y) = f\left(x(y)\right)\left|\frac{dx}{dy}\right| = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{y}e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2} \tag{5.24}$$

In figures 5.3, we plot some examples of the log-normal distribution for various values of the parameters μ and σ^2.

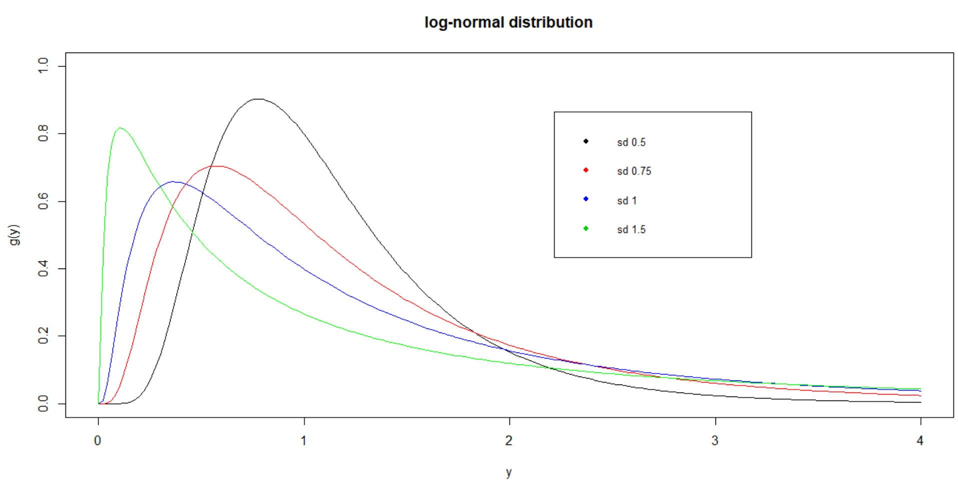

Figure 5.3a The PDF g(y) for the *log-normal distribution* for various values of the parameters σ and $\mu = 0$.

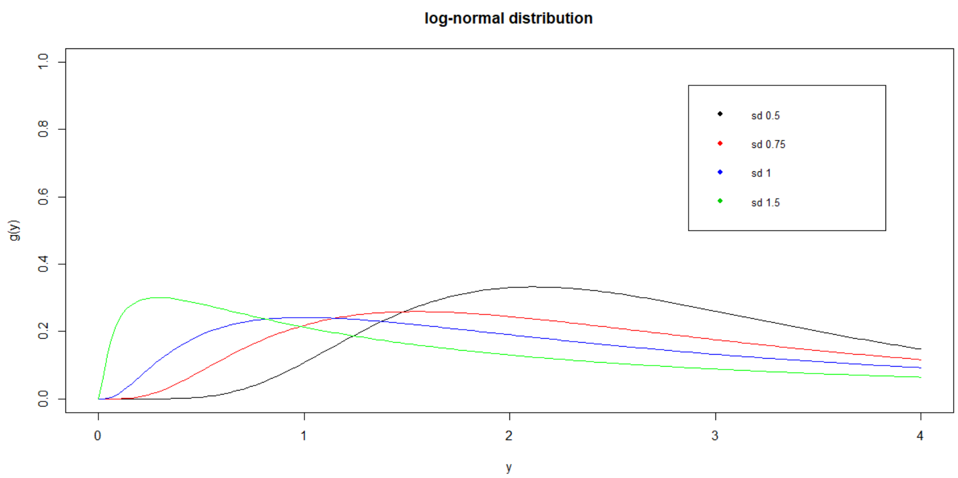

Figure 5.3b The PDF g(y) for various values of the parameters σ and $\mu = 1$.

We note that μ and σ^2 are not the mean and variance of the log-normal distribution, but rather the parameters of the corresponding Gaussian distribution for *X*. The *mean* and *variance* of *Y*, however, can be found using the MGF of *X*, which reads

$$M_X(t) = E\left[e^{tX}\right] = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \quad (5.25)$$

Thus, the *mean* and the *variance* of *Y* are

$$E[Y] = E\left[e^X\right] = M_X(1) = e^{\mu + \frac{1}{2}\sigma^2} \quad (5.26)$$

$$V[Y] = E\left[Y^2\right] - \left(E[Y]\right)^2 = E\left[e^{2X}\right] - \left(E\left[e^X\right]\right)^2$$

$$= M_X(2) - \left(M_X(1)\right)^2 = e^{2\mu + 2\sigma^2} - \left(e^{\mu + \frac{1}{2}\sigma^2}\right)^2 = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} = e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right)$$

5.3 The exponential and gamma distributions

The exponential distribution with positive parameter λ is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases} \quad (5.27)$$

and satisfies $\int_{-\infty}^{\infty} f(x)\,dx = 1$ as required. The *exponential distribution* occurs naturally if we consider the distribution of *the length of intervals between successive events in a Poisson process* or, equivalently, the distribution of the interval (i.e. the waiting time) before the first event. If the average number of events per unit interval is λ then on average there are λx events in interval *x*, so that from the Poisson distribution the probability that there will be no events in this interval is given by

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!} \text{ with } \lambda \to \lambda x \qquad \Rightarrow \qquad \Pr\left(\text{no events in interval x}\right) = e^{-\lambda x}$$

The probability that an event occurs in the next infinitestimal interval $[x, x+dx]$ is given by λdx, so that

$$\Pr\left(\text{the first event occurs in the interval } [x, x+dx]\right) = e^{-\lambda x}\lambda dx$$

Hence the required probability density function is given by

$$f(x) = \lambda e^{-\lambda x}$$

The *expectation* and *variance* of the exponential distribution can be evaluated as $1/\lambda$ and $(1/\lambda)^2$ respectively. The MGF is given by

$$M(t) = \frac{\lambda}{\lambda - t} \qquad (5.28)$$

Indeed,

$$M_X(t) = E\left[e^{tX}\right] = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}dx = \lambda\int_0^{\infty} e^{(t-\lambda)x}dx = \frac{\lambda}{\lambda - t}$$

$$M_X'(t) = \frac{\lambda}{(\lambda - t)^2} \quad \Rightarrow \quad E[X] = M_X'(0) = \frac{1}{\lambda}$$

$$M_X''(t) = \frac{2\lambda}{(\lambda - t)^3} \quad \Rightarrow \quad E\left[X^2\right] = M_X''(0) = \frac{2}{\lambda^2}$$

$$V[X] = E\left[X^2\right] - \left(E[X]\right)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

We may generalise the above discussion to obtain the PDF for *the interval between every rth event in a Poisson process* or, equivalently, the interval (waiting time) before the *r*th event. We begin by using the Poisson distribution to give

$$f(x)=e^{-\lambda}\frac{\lambda^{x}}{x!} \quad \text{with } \lambda\to\lambda x \qquad \Pr\left(r-1 \text{ events occur in interval } x\right)=e^{-\lambda x}\frac{(\lambda x)^{r-1}}{(r-1)!}\ ,$$

from which we obtain

$$\Pr\left(r\text{th event occurs in the interval } [x,x+dx]\right)=e^{-\lambda x}\frac{(\lambda x)^{r-1}}{(r-1)!}\lambda dx$$

Thus the required PDF is

$$f(x)=\frac{\lambda}{(r-1)!}(\lambda x)^{r-1}e^{-\lambda x} \qquad (5.29)$$

which is known as the *gamma distribution* of order *r* with parameter λ. Although our derivation applies only when *r* is a positive integer, the gamma distribution is defined for all positive *r* by replacing $(r-1)!$ by $\Gamma(r)$. If a random variable *X* is described by a *gamma distribution* of order *r* with parameter λ, we write $X\sim\gamma(\lambda,r)$; we note that the *exponential distribution* is the special case $\gamma(\lambda,1)$. The gamma distribution $\gamma(\lambda,r)$ is plotted in figure 5.4 for $\lambda=1$ and $r=1,2,5,10$. For large *r*, the gamma distribution tends to the Gaussian distribution whose mean and variance are specified by (5.31) below.

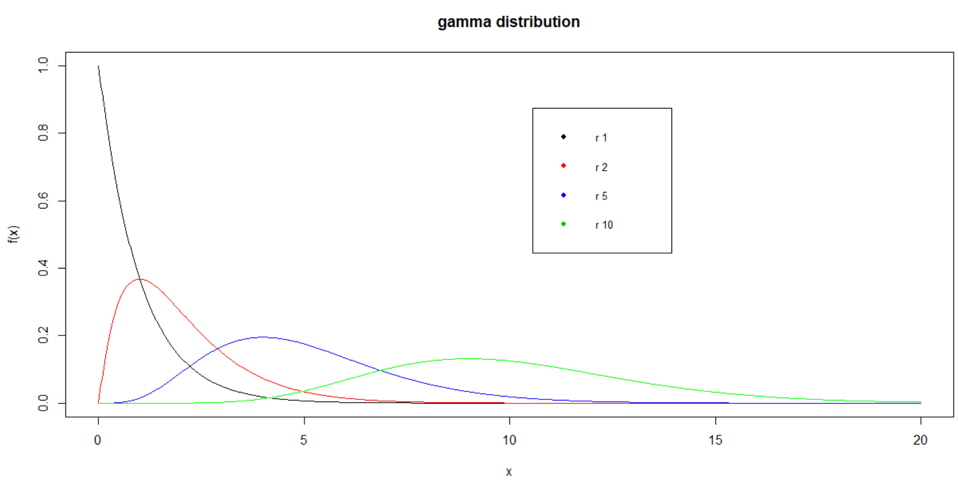

Figure 5.4 $f(x)$ for the gamma distributions $\gamma(\lambda,r)$ with $\lambda=1$ and $r=1,2,5,10$.

The MGF for the *gamma distribution* is obtained from that for the exponential distribution, by noting that we may consider the interval between every *r*th event in a Poisson process as the sum of *r* intervals between successive events. Thus the *r*th-order gamma variate is the sum of *r* independent exponentially distributed random variables. From (5.28), the MGF of the gamma distribution is given by

$$M(t) = \left(\frac{\lambda}{\lambda - t}\right)^r \tag{5.30}$$

from which the *mean* and *variance* are found to be

$$M'_X(t) = r\left(\frac{\lambda}{\lambda - t}\right)^{r-1} \frac{\lambda}{(\lambda - t)^2} \quad \Rightarrow \quad E[X] = M'_X(0) = \frac{r}{\lambda}$$

$$M''_X(t) = r(r-1)\left(\frac{\lambda}{\lambda - t}\right)^{r-2} \frac{\lambda^2}{(\lambda - t)^4} + r\left(\frac{\lambda}{\lambda - t}\right)^{r-1} \frac{2\lambda}{(\lambda - t)^3}$$

$$E\left[X^2\right] = M''_X(0) = \frac{r(r-1)}{\lambda^2} + \frac{2r}{\lambda^2} = \frac{r^2 + r}{\lambda^2}$$

$$V[X] = E\left[X^2\right] - \left(E[X]\right)^2 = \frac{r^2 + r}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}$$

$$E[X] = \frac{r}{\lambda} \qquad V[X] = \frac{r}{\lambda^2} \tag{5.31}$$

Multiple gamma distributions If $X_i \sim \gamma(\lambda, r_i)$, $i = 1, 2, \ldots, n$ are independent gamma variates then the random variable $Y = X_1 + X_2 + \ldots + X_n$ has MGF

$$M(t) = \prod_{i=1}^{n} \left(\frac{\lambda}{\lambda - t}\right)^{r_i} = \left(\frac{\lambda}{\lambda - t}\right)^{r_1 + r_2 + \ldots + r_n} \tag{5.32}$$

Thus *Y* is also a gamma variate, distributed as $Y \sim \gamma(\lambda, r_1 + r_2 + \ldots + r_n)$.

5.4 The chi-squared distribution

In section *Functions of random variables*, we showed that if *X* is Gaussian distributed with mean μ and variance σ^2, such that $X \sim N(\mu, \sigma^2)$, then the random variable $Y = (X - \mu)^2 / \sigma^2$ is distributed as the gamma distribution $Y \sim \gamma\left(\frac{1}{2}, \frac{1}{2}\right)$. Let us now consider *n* independent Gaussian random variables $X_i \sim N\left(\mu_i, \sigma_i^2\right), i = 1, 2, \ldots, n$, and define the new variable

$$\chi_n^2 = \sum_{i=1}^{n} \frac{\left(X_i - \mu_i\right)^2}{\sigma_i^2} \qquad (5.33)$$

Using the result (5.32) for *multiple gamma distributions*, χ_n^2 must be distributed as the gamma variate $\chi_n^2 \sim \gamma\left(\frac{1}{2}, \frac{1}{2}n\right)$, which from (5.29) has the PDF

$$f(x) = \frac{\lambda}{(r-1)!}(\lambda x)^{r-1} e^{-\lambda x} \qquad \Rightarrow \qquad f(x) = \frac{\frac{1}{2}}{\Gamma\left(\frac{1}{2}n\right)}\left(\frac{1}{2}x\right)^{\frac{n}{2}-1} e^{-\frac{1}{2}x}$$

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{1}{2}n\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \qquad (5.34)$$

This is known as the *chi-squared distribution* of order *n* and has numerous applications in statistics. Setting $\lambda = \frac{1}{2}$ and $r = \frac{1}{2}n$ in (5.31), we find that

$$E\left[\chi_n^2\right] = \frac{r}{\lambda} = n \ , \qquad V\left[\chi_n^2\right] = \frac{r}{\lambda^2} = 2n \qquad (5.35)$$

Generalisation: when the *n* Gaussian variables X_i are not linearly independent but are instead required to satisfy a linear constraint of the form

$$c_1 X_1 + c_2 X_2 + \ldots + c_n X_n = 0 \qquad (5.36)$$

in which the constants c_i are not all zero. In this case, it may be shown that the variable χ_n^2 defined in (5.33) is still described by a *chi squared distribution*, but one of order $n-1$. This result may be extended to show that if the *n* Gaussian variables X_i satisfy *m* linear constraints of the form (5.35) then the variable χ_n^2 defined in (5.33) is described by a *chi-squared distribution* of order $n-m$.

5.5 The uniform distribution

Finally we mention the very simple, but common, *uniform distribution*, which describes a continuous random variable that has a constant PDF over its allowed range of values. If the limits on *X* are *a* and *b* then

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{otherwise} \end{cases} \tag{5.36}$$

The MGF of the uniform distribution is found to be

$$M(t) = E\left[e^{tX}\right] = \int_a^b e^{tx} \frac{1}{b-a} dx = \frac{1}{b-a} \frac{e^{tx}}{t}\Bigg|_a^b$$

$$M(t) = \frac{e^{bt} - e^{at}}{(b-a)t} \tag{5.37}$$

and its mean and variance are given by

$$E[X] = \frac{a+b}{2} \qquad V[X] = \frac{(b-a)^2}{12} \tag{5.38}$$