standard deviation in density would exceed one-thousandth of the mean density? Assume that the number of particles in a given volume V is a random variable with a Poisson distribution; that is,

$$p_k = \frac{(\mu V)^k}{k!} e^{-\mu V}$$

is the probability of finding exactly k particles in the volume V, where there are μ particles per unit volume on the average. For air at 68°F and one atmosphere pressure, $\mu = 2.7 \times 10^{19}$ particles per cubic centimeter.

ANSWER. The sample would be a cube with side equal to 3.3×10^{-5} cm.

10.6 Common probability density functions

Uniform density function. The simplest density function for a random scalar is the uniform distribution:

$$p(x) = \begin{cases} \dfrac{1}{c}, & b - \dfrac{c}{2} \le x \le b + \dfrac{c}{2}, \\ 0, & x > b + \dfrac{c}{2}, x < b - \dfrac{c}{2}. \end{cases} \tag{10.6.1}$$

Obviously, we have

$$\int_{-\infty}^{\infty} p(x)\,dx = 1, \tag{10.6.2}$$

$$E(x) = b, \tag{10.6.3}$$

$$E(x - b)^2 = c^2/12. \tag{10.6.4}$$

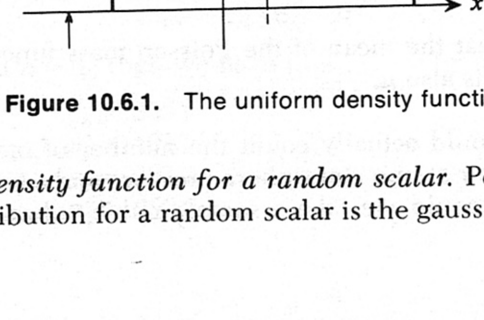


Figure 10.6.1. The uniform density function.

Gaussian density function for a random scalar. Perhaps the most common distribution for a random scalar is the gaussian distribution:

$$p(x) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2}\right]. \tag{10.6.5}$$

shown that

$$\int_{-\infty}^{\infty} p(x)\,dx = 1\,, \tag{10.6.6}$$

$$E(x) = \bar{x}\,, \tag{10.6.7}$$

$$E(x-\bar{x})^2 = \sigma^2\,. \tag{10.6.8}$$

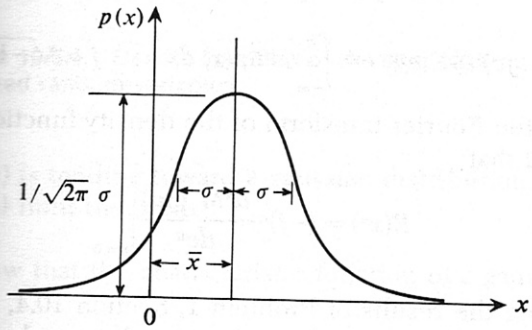


Figure 10.6.2. The gaussian density function.

ation for representing many complicated phenomena by density functions lies in the *central limit theorem*,† which , if x is the sum of N independent random quantities having density functions, then x tends to have a gaussian density as $N \to \infty$ (see Problems 1 and 2). The probability that x lies $\bar{x} - \xi$ and $\bar{x} + \xi$ is given by

$$\int_{\bar{x}-\xi}^{\bar{x}+\xi} p(x)\,dx = \frac{1}{(2\pi)^{1/2}\sigma}\int_{-\xi}^{\xi} e^{-(t^2/2\sigma^2)}\,dt$$

$$= \frac{2}{\sqrt{\pi}}\int_0^{\xi/\sqrt{2}\sigma} e^{-n^2}\,dn \triangleq \operatorname{erf}(\xi/\sqrt{2}\sigma)\,. \tag{10.6.9}$$

of this "normal probability integral" or "error function" are many places. Of particular interest are the values for $\xi = \sigma$, 3σ, given below:

---

ple, Cramer (1946), p. 317.

ξ	Value
σ	.683
2σ	.955
3σ	.997

The "three sigma" (3σ) value is often used in practical problems as virtually the upper bound on the variation from the mean, since the probability that x lies between -3σ and $+3\sigma$ is .997. Analogous to the concept of a generating function for the mass function, a *characteristic function* for the density function of a random variable is defined by

$$M_x(jv) \triangleq E(e^{jvx}) = \int_{-\infty}^{\infty} e^{jvx} p(x)\, dx\,, \qquad j = \sqrt{-1}\,, \tag{10.6.10}$$

which is just the Fourier transform of the density function. It can be easily verified that

$$E(x^n) = (-j)^n \frac{d^n M_x(jv)}{dv^n}\bigg|_{v=0} \tag{10.6.11}$$

Problem 1. Using the results of Problem 1, Section 10.4, consider the case in which x_1 and x_2 are independent random scalars, each uniformly distributed on the interval $(-\frac{1}{2}, \frac{1}{2})$. With $y = x_1 + x_2$, show that

$$p(y) = \begin{cases} 1 - |y|\,, & |y| < 1\,, \\ 0\,, & |y| > 1\,. \end{cases}$$

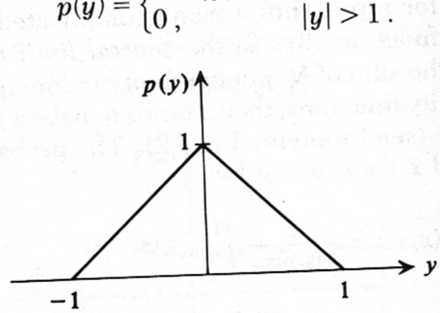


Figure 10.6.3. Density function of the sum of two uniformly distributed random variables.

Problem 2. Using the results of Problem 1 (and Problem 1 of Section 10.4 again), consider the case in which x_1, x_2, and x_3 are independent random scalars, each uniformly distributed on the interval $(-\frac{1}{2}, \frac{1}{2})$. With $y = x_1 + x_2 + x_3$, show that

$$p(y) = \begin{cases} \frac{3}{4} - y^2, & 0 \le |y| \le \frac{1}{2}, \\ \frac{1}{2}(\frac{3}{2} - |y|)^2, & \frac{1}{2} \le |y| \le \frac{3}{2}, \\ 0, & |y| \ge \frac{3}{2}. \end{cases}$$

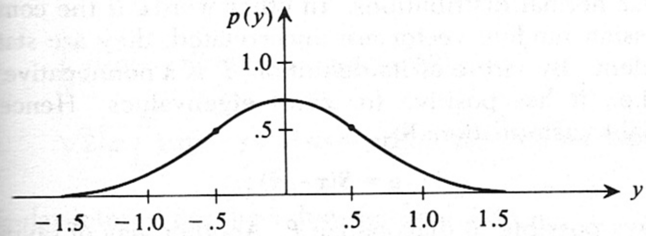


Figure 10.6.4. Density function of the sum of three uniformly distributed random variables.

Note that $p(y)$ is tending toward a gaussian distribution, as indicated by the central limit theorem.

Problem 3. Show that the characteristic function of a gaussian random variable is

$$M_x(jv) = \exp\left[jv\bar{x} - \frac{v^2\sigma^2}{2}\right].$$

10.7 Gaussian density function for a random vector

If x is a random n-vector, where the components can take on a continuous set of values, the most common probability density encountered in practice, and certainly the most important for this book, is the gaussian or normal distribution:†

$$p(x) = \frac{1}{(2\pi)^{n/2}|P|^{1/2}} \exp\left[-\frac{1}{2}(x - \bar{x})^T P^{-1}(x - \bar{x})\right]. \tag{10.7.1}$$

It can be shown that‡

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x)\, dx_1 \cdots dx_n = 1, \tag{10.7.2}$$

$$E(x) = \bar{x} = \text{mean value of vector}, \tag{10.7.3}$$

$$E[(x - \bar{x})(x - \bar{x})^T] = P = \text{covariance matrix of vector}, \tag{10.7.4}$$

† A convenient shorthand notation for this is "x is $N(\bar{x},P)$."

‡ Raiffa and Schlaifer (1961), pp. 246–251.

where $|P|$ is the determinant of P, P^{-1} is the matrix inverse of P. Note that $p(x)$ is completely characterized by giving only $\bar{x}$ and P.

If P is a diagonal matrix, then $x - \bar{x}$ has components that are statistically independent, since $p(x)$ may then be factored into a product of n scalar normal distributions. In other words, if the components of a gaussian random vector are uncorrelated, they are statistically independent. By virtue of its definition, P is a nonnegative definite matrix; i.e., it has positive (or zero) eigenvalues. Hence, by an orthogonal transformation, S,

$$y = S(x - \bar{x}), \tag{10.7.5}$$

it is always possible to diagonalize P. Another way of saying this is that the hypersurfaces of constant *likelihood* (constant values of probability density) in the x-space are hyperellipsoids, and, by a rotation of axes, it is possible to use the principal axes of these hyperellipsoids as coordinate axes.

We are often interested in the probability that x lies inside the hyperellipsoid:

$$(x - \bar{x})^T P^{-1} (x - \bar{x}) = l^2, \tag{10.7.6}$$

where l is a constant. By transforming to principal axes, this expression becomes

$$\frac{y_1^2}{\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} + \cdots + \frac{y_n^2}{\sigma_n^2} = l^2. \tag{10.7.7}$$

By another transformation, $z_i = (y_i/\sigma_i)$, this expression becomes the equation for a hypersphere in n dimensions:

$$z_1^2 + z_2^2 + \cdots + z_n^2 = r^2. \tag{10.7.8}$$

The probability of finding z inside this hypersphere is

$$\int\int_V \cdots \int \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}[z_1^2 + \cdots + z_n^2]\right\} dz_1, \ldots, dz_n, \tag{10.7.9}$$

where the integration is carried out over the volume V of the hypersphere, r, where

$$r^2 = z_1^2 + z_2^2 + \cdots + z_n^2. \tag{10.7.10}$$

In the z space $|P| = 1$, since all the variances are unity and all covariances are zero. Thus the probability of finding x inside the hyperellipsoid $(x - \bar{x})^T P^{-1} (x - \bar{x}) = l^2$ is

$$\left[\frac{1}{(2\pi)^{n/2}}\right] \int_0^l \exp\left(-\frac{1}{2}r^2\right) f(r)\, dr, \tag{10.7.11}$$

t of P, P^{-1} is the
haracterized by giving
en $x - \bar{x}$ has components
x) may then be factored
ns. In other words, if the
are uncorrelated, they are
s definition, P is a non
(or zero) eigenvalues.

$= S(x - \bar{x})$,

nalize P. Another way of
stant *likelihood* (constant
e are hyperellipsoids, and
e principal axes of these

n the probability that x lies

$)^T P^{-1}(x - \bar{x}) = l^2$,

nsforming to principal axes,

$+ \cdots + \frac{y_n^2}{\sigma_n^2} = l^2$.

$z_i = (y_i/\sigma_i)$, this expression
in n dimensions:

$z_2^2 + \cdots + z_n^2 = r^2$.

inside this hypersphere is

$\frac{1}{2}[z_1^2 + \cdots + z_n^2]\Big\} dz_1 \ldots$

ried out over the volume

$z_1^2 + z_2^2 + \cdots + z_n^2$.

ce all the variances are unity
e probability of finding x inside
$= l^2$ is

$\int_0^l \exp\left(-\frac{1}{2}r^2\right) f(r)\, dr$,

$f(r)\,dr$ is the spherically symmetric volume element in an n-dimensional space. For $n = 1,2,3$, this probability is given by

$$\sqrt{2/\pi}\int_0^l \exp(-\tfrac{1}{2}r^2)\,dr = \operatorname{erf}(l/\sqrt{2}),$$

$$\int_0^l \exp(-\tfrac{1}{2}r^2)\,r\,dr = 1 - \exp(-\tfrac{1}{2}l^2), \tag{10.7.12}$$

$$\sqrt{2/\pi}\int_0^l \exp(-\tfrac{1}{2}r^2)\,r^2\,dr = \operatorname{erf}(l/\sqrt{2}) - \sqrt{2/\pi}\,l\exp(-\tfrac{1}{2}l^2).$$

icular interest are the values for $l = 1,2,3$:

n/l	1	2	3
1	.683	.955	.997
2	.394	.865	.989
3	.200	.739	.971

se are often called the one-, two-, or three-sigma probabilities.

Consider a normally distributed two-dimensional vector with

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}.$$

nvalues of this covariance matrix are given by

$$\begin{vmatrix} 4-\sigma^2 & 1 \\ 1 & 1-\sigma^2 \end{vmatrix} = 0$$

$$\sigma^4 - 5\sigma^2 + 3 = 0, \quad \Rightarrow \sigma_1^2 = 4.30, \quad \sigma_2^2 = .70,$$

e eigenvectors are proportional to

$$\begin{bmatrix} 1 \\ .30 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -3.30 \end{bmatrix}.$$

lihood ellipses

$$(x_1, x_2)\begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = l^2$$

n in Figure 10.7.1 for $l = 1,2,3$. The probability of finding x the $l = 1$ ellipse is .394, inside the $l = 2$ ellipse is .865, and the $l = 3$ ellipse is .989.

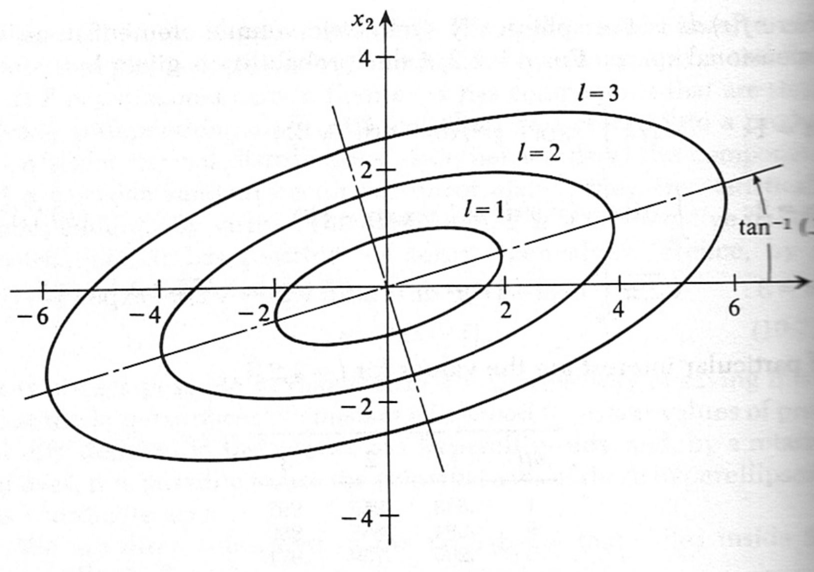


Figure 10.7.1. Likelihood ellipses for the example of two dimensional gaussian random vectors.

An important property of gaussian random vectors. The remaining part of this book depends heavily on one important property of gaussian random vectors; that is, *a linear combination of gaussian random vectors is also a gaussian random vector.* Stated analytically, if x is a gaussian random vector with mean $\bar{x}$ and covariance P_x, and $y = Ax + b$ where A is a constant matrix and b is a constant vector, then y is a gaussian random vector with mean $\bar{y}$ and covariance P_y where

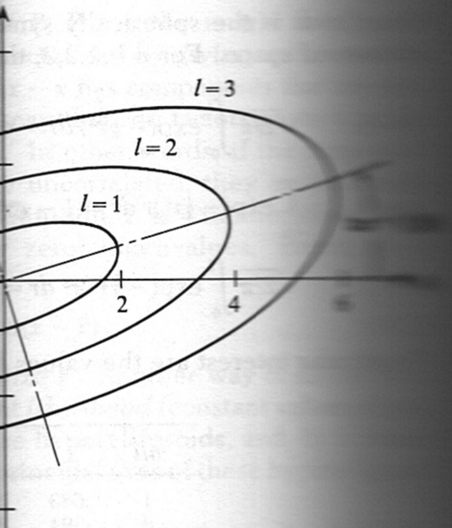
$$\bar{y} = A\bar{x} + b\,, \tag{10.7.13}$$

$$P_y = AP_xA^T\,. \tag{10.7.14}$$

The relations (10.7.13) and (10.7.14) follow very simply from the definition of expected values:

$$\begin{aligned}\bar{y} = E(y) &= \int_{-\infty}^{\infty} \cdots \int (Ax + b)p(x)\,dx_1 \cdots dx_n \\ &= A\int_{-\infty}^{\infty} \cdots \int xp(x)\,dx_1, \ldots, dx_n + b\int_{-\infty}^{\infty} \cdots \int p(x)\,dx_1 \cdots dx_n \\ &= A\bar{x} + b\,,\end{aligned}$$

and



ellipses for the example of
vectors.

ussian random vectors.
eavily on one important
is, *a linear combination*
an random vector. Stated
r with mean $\bar{x}$ and covariance
tant matrix and b is a constant
ector with mean $\bar{y}$ and covariance

$$= A\bar{x} + b\,,$$

$$= AP_xA^T\,.$$

10.7.14) follow very simply

$$x)\,dx_1 \cdots dx_n$$

$$, \ldots, dx_n + b\int_{-\infty}^{\infty} \cdots \int$$

$$= E[(y - \bar{y})(y - \bar{y})^T] = \int_{-\infty}^{\infty} \cdots \int A(x - \bar{x})(x - \bar{x})^T A^T p(x)\,dx_1 \cdots dx_n$$

$$= AP_xA^T\,.$$

show that y is a gaussian random vector is also quite simple if A nonsingular matrix.† The probability that y lies in a certain region the y-space, R_y, is equal to the probability that x lies in the corresponding region R_x of the x-space, that is, we have

$$\int_{R_y} \cdots \int p(y)\,dy_1 \cdots dy_n = \int_{R_x} \cdots \int p(x)\,dx_1 \cdots dx_n\,. \quad (10.7.15)$$

Changing variables of integration in the right-hand integral, using

$$dx_1 \cdots dx_n = |AA^T|^{-1/2}\,dy_1 \cdots dy_n\,, \quad (10.7.16)$$

the result

$$= |AA^T|^{-1/2}\,p(x) = \frac{|(AA^T)|^{-1/2}}{(2\pi)^{n/2}|P_x|^{1/2}} \times$$

$$\exp\left\{-\frac{1}{2}(y - \bar{y})^T(A^{-1})^T P_x^{-1} A^{-1}(y - \bar{y})\right\} \qquad \text{with} \qquad x = A^{-1}(y - b)$$

$$p(y) = \frac{1}{(2\pi)^{n/2}|P_y|^{1/2}} \exp\left\{-\frac{1}{2}(y - \bar{y})^T P_y^{-1}(y - \bar{y})\right\}, \quad (10.7.17)$$

which was to be shown.

1. Define the joint characteristic function of a random vector analogous to (10.6.10) via the use of multidimensional Fourier transform and show that, for gaussian x,

$$M_x(jv) = \exp(jv^T\bar{x} - \tfrac{1}{2}v^TPv)\,.$$

(See Davenport and Root, *Introduction to Random Signals and Noise,* McGraw-Hill, 1958, p. 153.)

2. Prove (10.7.13) and (10.7.14) for arbitrary A by using the result of Problem 1 (see Cramer (1946), p. 312).

3. If b is a gaussian random vector independent of x, with mean $\bar{b}$ and covariance P_b, show that Equations (10.7.13) and (10.7.14) modified to

† proof of the case in which A is singular, try Problem 2 or see Cramer (1946).