

10 — BIVARIATE DISTRIBUTIONS

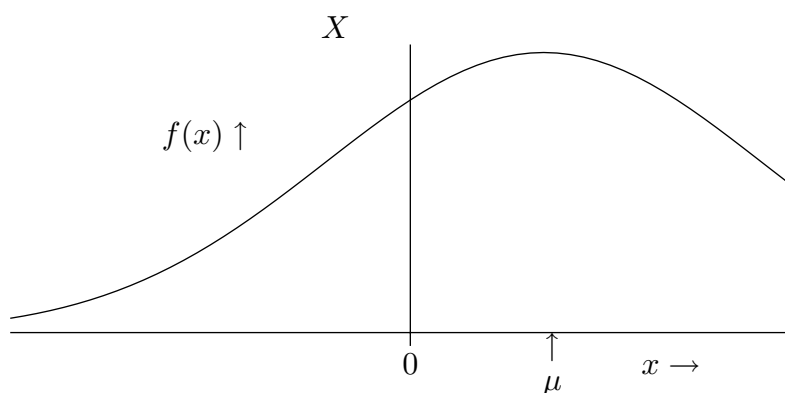
After some discussion of the Normal distribution, consideration is given to handling two continuous random variables.

The Normal Distribution

The probability density function $f(x)$ associated with the general Normal distribution is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (10.1)$$

The range of the Normal distribution is $-\infty$ to $+\infty$ and it will be shown that the total area under the curve is 1. It will also be shown that μ is the mean and that σ^2 is the variance. A graphical representation of the Normal distribution is:



It is immediately clear from (10.1) that $f(x)$ is symmetrical about $x = \mu$. This value of x is marked. When $\mu = 0$ the curve is symmetrical about the vertical axis. The value of σ^2 governs the width of the hump and it turns out that the inflexion points (one each side of the peak) are at $x = \mu \pm \sigma$ where $\sigma = \sqrt{\sigma^2}$ is the standard deviation.

When integrating expressions which incorporate the probability density function, it is essential to know the following standard result (a derivation is presented on page 10.10):

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi} \quad (10.2)$$

This result makes it easy to check that the total area under the curve is 1:

$$\text{Total Area} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $t = \frac{x - \mu}{\sqrt{2\sigma^2}}$ so $dx = \sqrt{2\sigma^2} dt$. Then:

$$\text{Total Area} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-t^2} \sqrt{2\sigma^2} dt = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1$$

From the symmetry of the probability density function (10.1) it is clear that the expectation $E(X) = \mu$ but to give a further illustration of the use of the standard result (10.2) the expectation is here evaluated by integration:

$$E(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x.e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $t = \frac{x - \mu}{\sqrt{2\sigma^2}}$ so $x = \sqrt{2\sigma^2} t + \mu$ and $dx = \sqrt{2\sigma^2} dt$. Then:

$$\begin{aligned} E(X) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (\sqrt{2\sigma^2}t + \mu).e^{-t^2} \sqrt{2\sigma^2} dt \\ &= \sqrt{\frac{2\sigma^2}{\pi}} \int_{-\infty}^{+\infty} t.e^{-t^2} dt + \frac{\mu}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt \\ &= \sqrt{\frac{2\sigma^2}{\pi}} \left[-\frac{1}{2}e^{-t^2} \right]_{-\infty}^{+\infty} + \frac{\mu}{\sqrt{\pi}} \cdot \sqrt{\pi} \\ &= \sqrt{\frac{2\sigma^2}{\pi}} [0 - 0] + \mu \\ &= \mu \end{aligned}$$

The variance, for once, is conveniently evaluated from the formula $V(X) = E((X - \mu)^2)$:

$$V(X) = E((X - \mu)^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x - \mu)^2 . e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $t = \frac{x - \mu}{\sqrt{2\sigma^2}}$ so $x - \mu = \sqrt{2\sigma^2} t$ and $dx = \sqrt{2\sigma^2} dt$. Then:

$$\begin{aligned} V(X) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} 2\sigma^2 t^2 . e^{-t^2} \sqrt{2\sigma^2} dt \\ &= -\frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} t \frac{d}{dt} (e^{-t^2}) dt \end{aligned}$$

The integrand has been put into a form ready for integration by parts:

$$\begin{aligned} V(X) &= -\frac{\sigma^2}{\sqrt{\pi}} \left[t.e^{-t^2} \right]_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt \\ &= -\frac{\sigma^2}{\sqrt{\pi}} [0 - 0] + \frac{\sigma^2}{\sqrt{\pi}} \cdot \sqrt{\pi} \\ &= \sigma^2 \end{aligned}$$

Standard Form of the Normal Distribution

The general Normal distribution is described as:

$$\text{Normal}(\mu, \sigma^2) \quad \text{or simply} \quad \text{N}(\mu, \sigma^2)$$

The smaller the variance σ^2 the narrower and taller the hump of the probability density function.

A particularly unfortunate difficulty with the Normal distribution is that integrating the probability density function between arbitrary limits is intractable. Thus, for arbitrary a and b , it is impossible to evaluate:

$$P(a \leq X < b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

The traditional way round this problem is to use tables though these days appropriate facilities are built into any decent spreadsheet application.

Unsurprisingly, tables are not available for a huge range of possible (μ, σ^2) pairs but the special case when the mean is zero and the variance is one, the distribution $\text{Normal}(0,1)$, is well documented. The probability distribution function for $\text{Normal}(0,1)$ is often referred to as $\Phi(x)$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

Tables of $\Phi(x)$ are readily available so, given a random variable X distributed $\text{Normal}(0,1)$, it is easy to determine the probability that X lies in the range a to b . This is:

$$P(a \leq X < b) = \Phi(b) - \Phi(a)$$

It is now straightforward to see how to deal with the general case of a random variable X distributed $\text{Normal}(\mu, \sigma^2)$. To determine the probability that X lies in the range a to b first reconsider:

$$P(a \leq X < b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $t = \frac{x - \mu}{\sigma}$ so $dx = \sigma dt$. Then:

$$P(a \leq X < b) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} \sigma dt = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt$$

The integration is thereby transformed into that for the distribution $\text{Normal}(0,1)$ and is said to be in *standard form*. Noting the new limits:

$$P(a \leq X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

The Central Limit Theorem

No course on probability or statistics is complete without at least a passing reference to the Central Limit Theorem. This is an extraordinarily powerful theorem but only the most token of introductory remarks will be made about it here.

Suppose X_1, X_2, \dots, X_n are n independent and identically distributed random variables each with mean μ and variance σ^2 . Consider two derived random variables Y and Z which are respectively the sum and mean of X_1, X_2, \dots, X_n :

$$Y = X_1 + X_2 + \dots + X_n \quad \text{and} \quad Z = \frac{X_1 + X_2 + \dots + X_n}{n}$$

From the rules for Expectation and Variance discussed on pages 4.6 and 4.8 it is simple to determine the Expectation and Variance of Y and Z :

$$E(Y) = n\mu, \quad V(Y) = n\sigma^2 \quad \text{and} \quad E(Z) = \mu, \quad V(Z) = \frac{\sigma^2}{n}$$

The Central Limit Theorem addresses the problem of how the derived random variables Y and Z are distributed and asserts that, as n increases indefinitely, both Y and Z tend to a Normal distribution *whatever* the distribution of the individual X_i . More specifically:

$$Y \text{ tends to } N(n\mu, n\sigma^2) \quad \text{and} \quad Z \text{ tends to } N\left(\mu, \frac{\sigma^2}{n}\right)$$

In many cases Y and Z are approximately Normal for remarkably small values of n .

Bivariate Distributions — Reference Discrete Example

It has been noted that many formulae for continuous distributions parallel equivalent formulae for discrete distributions. In introducing examples of two *continuous* random variables it is useful to employ a reference example of two *discrete* random variables.

Consider two discrete random variables X and Y whose values are r and s respectively and suppose that the probability of the event $\{X = r\} \cap \{Y = s\}$ is given by:

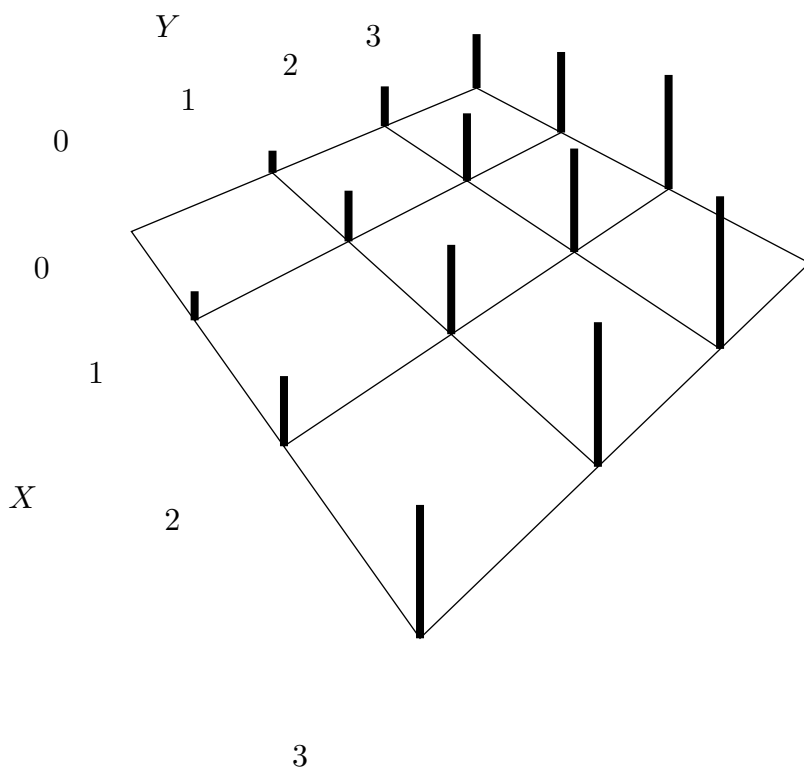
$$P(X = r, Y = s) = \begin{cases} \frac{r+s}{48}, & \text{if } 0 \leq r, s \leq 3 \\ 0, & \text{otherwise} \end{cases}$$

The probabilities may be tabulated thus:

		Y				
			s	→		
		0	1	2	3	
X	0	$\frac{0}{48}$	$\frac{1}{48}$	$\frac{2}{48}$	$\frac{3}{48}$	$\frac{6}{48}$
	1	$\frac{1}{48}$	$\frac{2}{48}$	$\frac{3}{48}$	$\frac{4}{48}$	$\frac{10}{48}$
r	2	$\frac{2}{48}$	$\frac{3}{48}$	$\frac{4}{48}$	$\frac{5}{48}$	$\frac{14}{48}$
	↓	$\frac{3}{48}$	$\frac{4}{48}$	$\frac{5}{48}$	$\frac{6}{48}$	$\frac{18}{48}$
		$\frac{6}{48}$	$\frac{10}{48}$	$\frac{14}{48}$	$\frac{18}{48}$	
			P(Y = s)	→		

$P(X = r)$
 \downarrow

The following is a graphical representation of the same function:



The probabilities exist for only integer values of r and s so the effect is of a kind of pegboard. Different lengths of peg correspond to different probabilities.

Axiom II requires the sum of all the entries in the table (which is the total length of the pegs) to be 1.

The table also includes the marginal sums which separately tabulate the probabilities $P(X = r)$ and $P(Y = s)$.

Bivariate Distributions — Continuous Random Variables

When there are two *continuous* random variables, the equivalent of the two-dimensional array is a region of the x - y (cartesian) plane. Above the plane, over the region of interest, is a surface which represents the probability density function associated with a *bivariate distribution*.

Suppose X and Y are two continuous random variables and that their values, x and y respectively, are constrained to lie within some region R of the cartesian plane. The associated probability density function has the general form $f_{XY}(x, y)$ and, regarding this function as defining a surface over region R , axiom II requires:

$$\iint_R f_{XY}(x, y) dx dy = 1$$

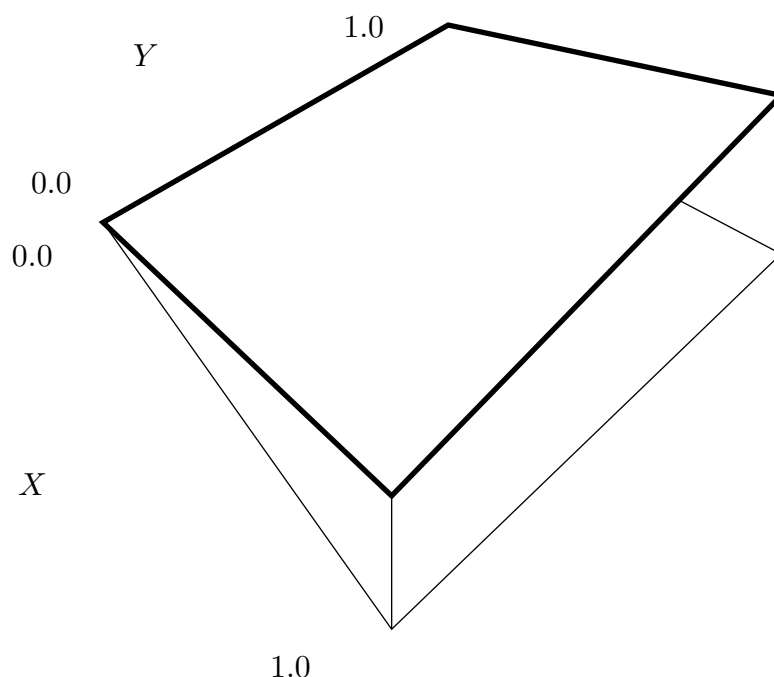
This is equivalent to requiring that the volume under the surface is 1 and corresponds to the requirement that the total length of the pegs in the pegboard is 1.

As an illustration, consider a continuous analogy of the reference example of two discrete random variables.

Suppose X and Y are two continuous random variables and that their values x and y are constrained to lie in the unit square $0 \leq x, y < 1$. Suppose further that the associated bivariate probability density function is:

$$f_{XY}(x, y) = \begin{cases} x + y, & \text{if } 0 \leq x, y < 1 \\ 0, & \text{otherwise} \end{cases}$$

This probability density function can be regarded as defining a surface over the unit square. Continuous functions cannot satisfactorily be tabulated but it is not difficult to depict a graphical representation of $f_{XY}(x, y)$:



Note that when $x = y = 1$ the value of $f_{XY}(x, y)$ is 2, an impossible value for a probability but a perfectly possible value for a probability density function. Probabilities correspond to volumes now and it is easy to check that the total volume under the surface is 1. Given that region R represents the specified unit square:

$$\begin{aligned} \iint_R f_{XY}(x, y) dx dy &= \int_0^1 \int_0^1 (x + y) dx dy = \int_0^1 \left[\frac{x^2}{2} + yx \right]_0^1 dy \\ &= \int_0^1 \left(\frac{1}{2} + y \right) dy = \left[\frac{y}{2} + \frac{y^2}{2} \right]_0^1 = 1 \end{aligned}$$

In the general case where the probability density function is $f_{XY}(x, y)$, the probability that (x, y) lies in a particular sub-region R_s of R is given by:

$$P(X, Y \text{ lies in } R_s) = \iint_{R_s} f_{XY}(x, y) dx dy$$

In the example case, suppose that the sub-region R_s is defined by the half-unit square $0 \leq x, y < \frac{1}{2}$:

$$\begin{aligned} P(X, Y \text{ lies in } R_s) &= \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} (x + y) dx dy = \int_0^{\frac{1}{2}} \left[\frac{x^2}{2} + yx \right]_0^{\frac{1}{2}} dy \\ &= \int_0^{\frac{1}{2}} \left(\frac{1}{8} + \frac{y}{2} \right) dy = \left[\frac{y}{8} + \frac{y^2}{4} \right]_0^{\frac{1}{2}} = \frac{1}{8} \end{aligned}$$

In the reference discrete example (where the values of both random variables are confined to the range 0 to 3) this result loosely corresponds to determining $P(X, Y < 1\frac{1}{2})$:

$$P(X, Y < 1\frac{1}{2}) = \sum_{r=0}^1 \sum_{s=0}^1 P(X = r, Y = s) = \frac{0}{48} + \frac{1}{48} + \frac{1}{48} + \frac{2}{48} = \frac{4}{48} = \frac{1}{12}$$

Why is the answer different?

The Equivalent of Marginal Sums

With two discrete random variables, the marginal sums $P(X = r)$ and $P(Y = s)$ are given by the relationships:

$$P(X = r) = \sum_s P(X = r, Y = s) \quad \text{and} \quad P(Y = s) = \sum_r P(X = r, Y = s)$$

A pair of continuous random variables X and Y governed by a bivariate distribution function $f_{XY}(x, y)$ will, separately, have associated probability density functions $f_X(x)$ and $f_Y(y)$. By analogy with the discrete case, these functions are given by the relationships:

$$f_X(x) = \int_{y_{\min}}^{y_{\max}} f_{XY}(x, y) dy \quad \text{and} \quad f_Y(y) = \int_{x_{\min}}^{x_{\max}} f_{XY}(x, y) dx$$

The limits of integration merit a moment's thought. The region, R , over which $f_{XY}(x, y)$ is defined is not, in general, a square. Accordingly, the valid range of y depends on x and, in consequence, the limits y_{\min} and y_{\max} will depend on x . Likewise, the limits x_{\min} and x_{\max} will depend on y .

In the example case, where R is the unit square, there is no difficulty about the limits:

$$f_X(x) = \int_0^1 (x + y) dy = \left[xy + \frac{y^2}{2} \right]_0^1 = x + \frac{1}{2}$$

and:

$$f_Y(y) = \int_0^1 (x + y) dx = \left[\frac{x^2}{2} + yx \right]_0^1 = y + \frac{1}{2}$$

The Equivalent of Marginal Sums and Independence

With two discrete random variables, the two sets of marginal sums each sum to 1. With continuous random variables, integrating each of the functions $f_X(x)$ and $f_Y(y)$ must likewise yield 1. It is easy to verify that this requirement is satisfied in the present case:

$$\int_0^1 \left(x + \frac{1}{2}\right) dx = \left[\frac{x^2}{2} + \frac{x}{2}\right]_0^1 = 1 \quad \text{and} \quad \int_0^1 \left(y + \frac{1}{2}\right) dy = \left[\frac{y^2}{2} + \frac{y}{2}\right]_0^1 = 1$$

With two discrete random variables, marginal sums are used to test for independence. Two variables X and Y are said to be independent if:

$$P(X = r, Y = s) = P(X = r) \cdot P(Y = s) \quad \text{for all } r, s$$

By analogy, two continuous random variables X and Y are said to be independent if:

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$$

In the present case:

$$f_{XY}(x, y) = x + y \quad \text{and} \quad f_X(x) \cdot f_Y(y) = \left(x + \frac{1}{2}\right) \left(y + \frac{1}{2}\right) = xy + \frac{x+y}{2} + \frac{1}{4}$$

Clearly two continuous random variables X and Y whose probability density function is $x + y$ are *not* independent but the function just derived can be dressed up as bivariate probability density function whose associated random variables *are* independent:

$$f_{XY}(x, y) = \begin{cases} (4xy + 2x + 2y + 1)/4 & \text{if } 0 \leq x, y < 1 \\ 0, & \text{otherwise} \end{cases}$$

Illustration — The Uniform Distribution

Suppose X and Y are independent and that both are distributed Uniform(0,1). Their associated probability density functions are:

$$f_X(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad f_Y(y) = \begin{cases} 1, & \text{if } 0 \leq y < 1 \\ 0, & \text{otherwise} \end{cases}$$

Given this trivial case, the product is clearly the bivariate distribution function:

$$f_{XY}(x, y) = \begin{cases} 1, & \text{if } 0 \leq x, y < 1 \\ 0, & \text{otherwise} \end{cases}$$

The surface defined by $f_{XY}(x, y)$ is a unit cube so it is readily seen that:

$$\iint_R f_{XY}(x, y) dx dy = 1$$

Illustration — The Normal Distribution

Suppose X and Y are independent and that both are distributed $\text{Normal}(0,1)$. Their associated probability density functions are:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad \text{and} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

In this case the product leads to the probability density function:

$$f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} \quad (10.3)$$

The surface approximates that obtained by placing a large ball in the centre of a table and then draping a tablecloth over the ball.

Glossary

The following technical terms have been introduced:

standard form

bivariate distribution

Exercises — X

1. Find the points of inflexion of the probability density function associated with a random variable which is distributed $\text{Normal}(\mu, \sigma^2)$. Hence find the points at which the tangents at the points of inflection intersect the x -axis.
2. The Cauchy distribution has probability density function:

$$f(x) = \frac{c}{1+x^2} \quad \text{where} \quad -\infty < x < +\infty$$

Evaluate c . Find the probability distribution function $F(x)$. Calculate $P(0 \leq x < 1)$.

3. Two continuous random variables X and Y have the following bivariate probability function which is defined over the unit square:

$$f_{XY}(x, y) = \begin{cases} (9 - 6x - 6y + 4xy)/4 & \text{if } 0 \leq x, y < 1 \\ 0, & \text{otherwise} \end{cases}$$

- (a) Given that R is the unit square, verify that $\iint_R f_{XY}(x, y) dx dy = 1$.
- (b) Determine $f_X(x)$ and $f_Y(y)$.
- (c) Hence state whether or not the two random variables are independent.

ADDENDUM — AN IMPORTANT INTEGRATION

For arbitrary limits a and b it is impossible to evaluate $\int_a^b e^{-x^2} dx$ but evaluation *is* possible if $a = -\infty$ and $b = +\infty$ and an informal analysis is presented below.

As a preliminary observation note that:

$$\int_0^\infty e^{-x^2} dx = \sqrt{\int_0^\infty e^{-x^2} dx \int_0^\infty e^{-y^2} dy} = \sqrt{\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dx dy}$$

The item under the square root sign can be integrated by the substitution of two new variables r and θ using the transformation functions:

$$x = r \cos \theta \quad \text{and} \quad y = r \sin \theta$$

Further discussion of integration by the substitution of two new variables will be given on pages 12.1 and 12.2. In the present case, the integration is transformed thus:

$$\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dx dy = \int_0^{\frac{\pi}{2}} \int_0^\infty e^{-r^2} r dr d\theta$$

The first pair of limits reflects integration over a square constituting the quadrant of the cartesian plane in which x and y are both positive. The second pair of limits reflects integration over a quarter circle in the same quadrant. Since the function being integrated is strongly convergent as x and y or r increase, it is valid to equate the two integrations.

The right-hand integration is straightforward:

$$\int_0^{\frac{\pi}{2}} \int_0^\infty e^{-r^2} r dr d\theta = \int_0^{\frac{\pi}{2}} \left[\frac{-e^{-r^2}}{2} \right]_0^\infty d\theta = \int_0^{\frac{\pi}{2}} \frac{1}{2} d\theta = \left[\frac{\theta}{2} \right]_0^{\frac{\pi}{2}} = \frac{\pi}{4}$$

Accordingly:

$$\int_0^\infty e^{-x^2} dx = \sqrt{\frac{\pi}{4}} = \frac{\sqrt{\pi}}{2}$$

This leads, by symmetry, to the standard result quoted as (10.2):

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$