11 — TRANSFORMING DENSITY FUNCTIONS

It can be expedient to use a *transformation function* to transform one probability density function into another. As an introduction to this topic, it is helpful to recapitulate the method of integration by substitution of a new variable.

Integration by Substitution of a new Variable

Imagine that a newcomer to integration comes across the following:

$$\int_0^{\sqrt{\frac{\pi}{2}}} 2x \cos x^2 \, dx$$

Assuming that the newcomer doesn't notice that the integrand is the derivative of $\sin x^2$, one way to proceed would be to substitute a new variable y for x^2:

Let $y = x^2$

Replace the limits $x = 0$ and $x = \sqrt{\frac{\pi}{2}}$ by $y = 0$ and $y = \frac{\pi}{2}$

Replace $2x \cos x^2$ by $2\sqrt{y}\, \cos y$

Note that $x = \sqrt{y}$ and hence $\frac{dx}{dy} = \frac{1}{2\sqrt{y}}$ and so replace dx by $\frac{dy}{2\sqrt{y}}$

The original problem is thereby transformed into the following integration:

$$\int_0^{\frac{\pi}{2}} \cos y \, dy = \Big[\sin y\Big]_0^{\frac{\pi}{2}} = 1$$

The General Case

It is instructive to develop the general case alongside the above example:

	General Case	*Above Example*
	$\int_a^b f(x)\, dx$	$\int_0^{\sqrt{\frac{\pi}{2}}} 2x \cos x^2 \, dx$
Choose a transformation function $y(x)$		$y(x) = x^2$
Note its inverse $x(y)$		$x(y) = \sqrt{y}$
Replace the limits by $y(a)$ and $y(b)$		0 and $\frac{\pi}{2}$
Replace $f(x)$ by $f\big(x(y)\big)$		$2\sqrt{y}\, \cos y$
Replace dx by $\dfrac{dx}{dy}\, dy$		$\dfrac{1}{2\sqrt{y}}\, dy$
Result is $\displaystyle\int_{y(a)}^{y(b)} f\big(x(y)\big) \frac{dx}{dy}\, dy$		$\int_0^{\frac{\pi}{2}} \cos y \, dy$

Application to Probability Density Functions

The previous section informally leads to the general formula for integration by substitution of a new variable:

$$\int_a^b f(x)\,dx = \int_{y(a)}^{y(b)} f\big(x(y)\big)\,\frac{dx}{dy}\,dy \tag{11.1}$$

This formula has direct application to the process of transforming probability density functions...

Suppose X is a random variable whose probability density function is $f(x)$.

By definition:

$$\mathrm{P}(a \leqslant X < b) = \int_a^b f(x)\,dx \tag{11.2}$$

Any function of a random variable is itself a random variable and, if y is taken as some transformation function, $y(X)$ will be a derived random variable. Let $Y = y(X)$.

Notice that if $X = a$ the derived random variable $Y = y(a)$ and if $X = b$, $Y = y(b)$. Moreover, (subject to certain assumptions about y) if $a \leqslant X < b$ then $y(a) \leqslant Y < y(b)$ and $\mathrm{P}\big(y(a) \leqslant Y < y(b)\big) = \mathrm{P}(a \leqslant X < b)$. Hence, by (11.2) and (11.1):

$$\mathrm{P}\big(y(a) \leqslant Y < y(b)\big) = \mathrm{P}(a \leqslant X < b) = \int_a^b f(x)\,dx = \int_{y(a)}^{y(b)} f\big(x(y)\big)\,\frac{dx}{dy}\,dy \tag{11.3}$$

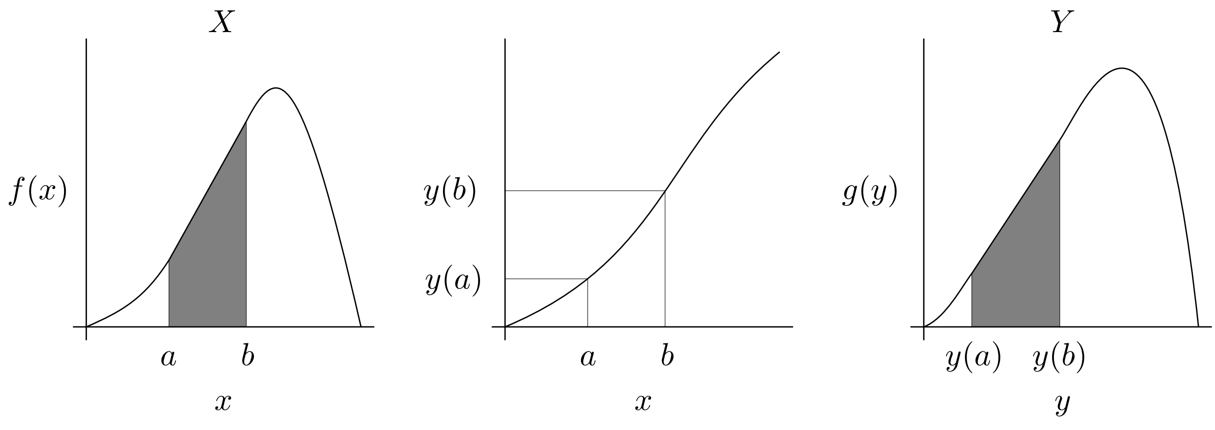
Notice that the right-hand integrand $f\big(x(y)\big)\,\dfrac{dx}{dy}$ is expressed wholly in terms of y.

Calling this integrand $g(y)$:

$$\mathrm{P}\big(y(a) \leqslant Y < y(b)\big) = \int_{y(a)}^{y(b)} g(y)\,dy$$

This demonstrates that $g(y)$ is the probability density function associated with Y.

The transformation is illustrated by the following figures in which the function $f(x)$ (on the left) is transformed by $y(x)$ (centre) into the new function $g(y)$ (right):



Observations and Constraints

The crucial step is (11.3). One imagines noting a sequence of values of a random variable X and for each value in the range a to b using a transformation function $y(x)$ to compute a value for a derived random variable Y.

Given certain assumptions about $y(x)$, the value of Y must be in the range $y(a)$ to $y(b)$ and the probability of Y being in this range is clearly the same as the probability of X being in the range a to b.

In summary: the shaded region in the right-hand figure has the same area as the shaded region in the left-hand figure.

There are three important conditions that any probability density function $f(x)$ has to satisfy:

- $f(x)$ must be single valued for all x
- $f(x) \geqslant 0$ for all x
- $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx = 1$

Often the function usefully applies over some finite interval of x and is deemed to be zero outside this interval. The function $2x\cos x^2$ could be used in the specification of a probability density function:

$$f(x) = \begin{cases} 2x\cos x^2, & \text{if } 0 \leqslant x < \sqrt{\frac{\pi}{2}} \\ 0, & \text{otherwise} \end{cases}$$

By inspection, $f(x)$ is single valued and non-negative and, given the analysis on page 11.1, the integral from $-\infty$ to $+\infty$ is one.

The constraints on the specification of a probability density function result in implicit constraints on any transformation function $y(x)$, most importantly:

- Throughout the useful range of x, both $y(x)$ and its inverse $x(y)$ must be defined and must be single-valued.
- Throughout this range, $\dfrac{dx}{dy}$ must be defined and *either* $\dfrac{dx}{dy} \geqslant 0$ *or* $\dfrac{dx}{dy} \leqslant 0$.

If $\frac{dx}{dy}$ were to change sign there would be values of x for which $y(x)$ would be multivalued (as would be the case if the graph of $y(x)$ were an S-shaped curve).

A consequence of the constraints is that any practical transformation function $y(x)$ must either increase monotonically over the useful range of x (in which case for any $a < b$, $y(a) < y(b)$) or decrease monotonically (in which case for any $a < b$, $y(a) > y(b)$).

Noting these constraints, it is customary for the relationship between a probability density function $f(x)$, the inverse $x(y)$ of a transformation function, and the derived probability density function $g(y)$ to be written:

$$g(y) = f\big(x(y)\big)\left|\frac{dx}{dy}\right| \tag{11.4}$$

Example I

Take a particular random variable X whose probability density function $f(x)$ is:

$$f(x) = \begin{cases} \dfrac{x}{2}, & \text{if } 0 \leqslant x < 2 \\ 0, & \text{otherwise} \end{cases}$$

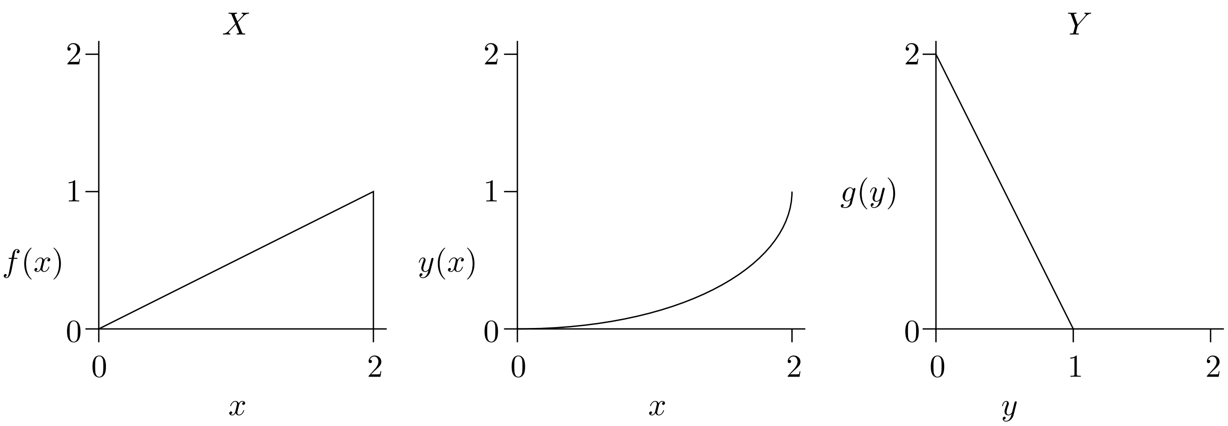
Suppose the transformation function $y(x)$ is:

$$y(x) = 1 - \frac{\sqrt{4 - x^2}}{2}$$

Note that the useful part of the range of x is 0 to 2 and, over this range, $y(x)$ *increases* monotonically from 0 to 1.

Let $Y = y(X)$, the derived random variable, and let $g(y)$ be the probability density function associated with Y. What is the function $g(y)$?

The problem is illustrated by the following figures:



First, derive $x(y)$ the inverse of the function $y(x)$.

Given:

$$y = 1 - \frac{\sqrt{4 - x^2}}{2}$$

$$4(y-1)^2 = 4 - x^2$$

So:

$$4y^2 - 8y + 4 = 4 - x^2$$

$$x^2 = 4y(2 - y)$$

$$x = 2\sqrt{y(2 - y)}$$

Accordingly:

$$f\big(x(y)\big) = \sqrt{y(2-y)} \qquad \text{and} \qquad \frac{dx}{dy} = \frac{2(1-y)}{\sqrt{y(2-y)}}$$

From (11.4):

$$g(y) = f\big(x(y)\big)\left|\frac{dx}{dy}\right| = 2(1-y)$$

As illustrated in the figures, the function $y(x)$ transforms one triangular distribution $f(x)$ into another $g(y)$. The two triangles are opposite ways round and the transformation function $y(x)$ has to ensure that although low values of X are relatively rare, low values of Y are common.

Expressing this informally: $y(x)$ stays low for most of the range of x so that even when x is well over one, the value of y is well under a half. This ensures that the transformation shifts the bias appropriately.

An Alternative Question

In the example, a probability density function and a transformation function were given and the requirement was to determine what new probability density function results.

Suppose instead that two probability density functions are given and the requirement is to find a function which transforms one into the other.

Take the particular functions used in the previous example and pose the question as follows. Given:

$$f(x) = \begin{cases} \dfrac{x}{2}, & \text{if } 0 \leqslant x < 2 \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad g(y) = \begin{cases} 2(1-y), & \text{if } 0 \leqslant y < 1 \\ 0, & \text{otherwise} \end{cases}$$

determine the function $y(x)$ which will transform $f(x)$ into $g(y)$.

From the relationship $g(y) = f\big(x(y)\big)\left|\dfrac{dx}{dy}\right|$:

$$2(1-y) = \frac{x}{2}\,\frac{dx}{dy}$$

or:

$$x\frac{dx}{dy} = 4(1-y)$$

This differential equation is readily solved and yields:

$$\frac{x^2}{2} = 4y - 2y^2 + c$$

Since $X = 0$ has to transform into $Y = 0$, the constant $c = 0$.

Continuing:

$$x^2 = 4(2y - y^2)$$

Hence the inverse function $x(y)$ is:

$$x(y) = 2\sqrt{y(2-y)}$$

A little more processing is required to determine $y(x)$:

$$y^2 - 2y + 1 = 1 - \frac{x^2}{4}$$

Hence:

$$(y-1)^2 = 1 - \frac{x^2}{4}$$

This leads to:

$$y = 1 \pm \sqrt{1 - \frac{x^2}{4}}$$

Choice of sign is important. Note, again, that $X = 0$ has to transform into $Y = 0$ and hence minus is appropriate.

This gives the solution:

$$y(x) = 1 - \frac{\sqrt{4 - x^2}}{2}$$

Transforming a Uniform Distribution

It would be unusual to wish to transform a triangular distribution but there is a good reason for wanting to be able to transform a uniform distribution into something else.

The generation of a uniform distribution by computer is a well-understood process and a typical programming language will be supplied with a library procedure to generate a random variable whose values are uniformly distributed.

All that remains to generate a random variable which is distributed differently is to use an appropriate transformation function.

It is very common to start with a distribution which is Uniform(0,1) which is to say that the probability density function $f(x)$ is:

$$f(x) = \begin{cases} 1, & \text{if } 0 \leqslant x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Over the useful range of x, the relationship $g(y) = f\big(x(y)\big)\left|\frac{dx}{dy}\right|$ simplifies to:

$$g(y) = \left|\frac{dx}{dy}\right| \tag{11.5}$$

Example II

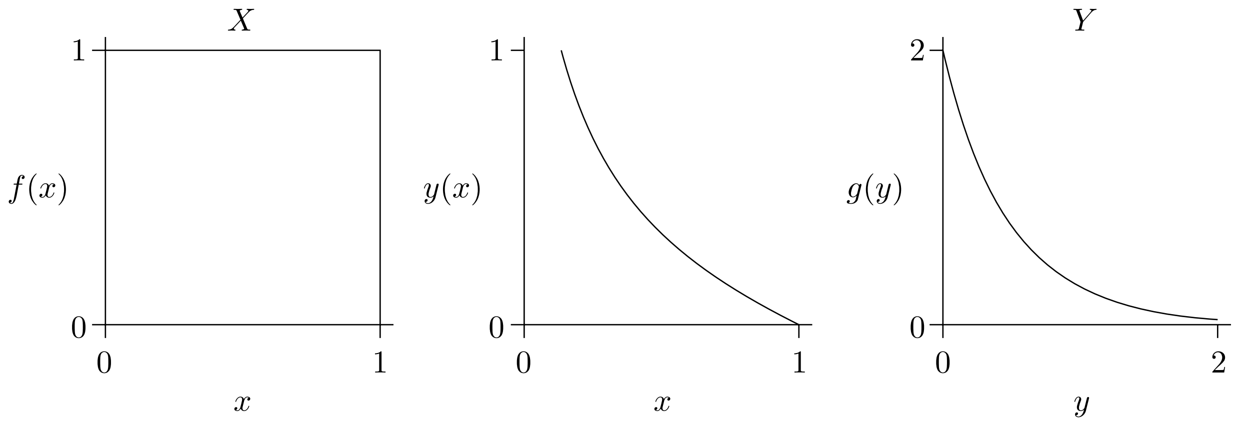
Take a random variable X whose probability density function $f(x)$ is Uniform(0,1) and suppose that the transformation function $y(x)$ is:

$$y(x) = -\frac{1}{\lambda}\ln x \qquad (\lambda > 0)$$

Note that the useful part of the range of x is 0 to 1 and, over this range, $y(x)$ *decreases* monotonically from ∞ to 0.

Let $Y = y(X)$ and let $g(y)$ be the probability density function associated with Y. What is the function $g(y)$?

The problem is illustrated by the following figures (in which $\lambda = 2$):



First, derive $x(y)$ the inverse of the function $y(x)$.

Given:

$$y = -\frac{1}{\lambda}\ln x$$

$$x = e^{-\lambda y}$$

Accordingly:

$$\frac{dx}{dy} = -\lambda.e^{-\lambda y}$$

Given that $\lambda > 0$ this derivative $\frac{dx}{dy}$ is everywhere negative.

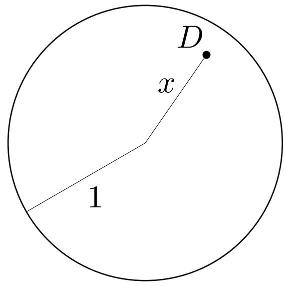
From (11.5):

$$g(y) = \left|\frac{dx}{dy}\right| = \lambda.e^{-\lambda y}$$

As illustrated in the figures, the function $y(x)$ transforms the distribution $f(x)$ which is Uniform(0,1) into $g(y)$ which is the exponential distribution.

Example III — Introduction

Suppose raindrops fall in a uniformly distributed way onto the surface of a circular pond which has unit radius.



Let X be a random variable whose value x is the distance of a raindrop (shown at D in the figure) from the centre of the pond. What is the probability density function $f(x)$ associated with X?

Consider a narrow annular concentric strip of radius x and width δx. The area of this strip is $2\pi x\,\delta x$. The area of the pond as a whole is $\pi.1^2$.

Hence:

$$\mathrm{P}(x \leqslant X < x + \delta x) = \frac{2\pi x\,\delta x}{\pi.1^2}$$

The probability density function $f(x)$ is therefore $2x$ or, more strictly:

$$f(x) = \begin{cases} 2x, & \text{if } 0 \leqslant x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Note, as a check, that $f(x)$ is single valued and non-negative and its integral from $-\infty$ to $+\infty$ is one.

This is another triangular distribution and leads to the unsurprising result that more raindrops fall close to the edge of the pond than fall close to the centre.

Example III — Transformation

The value of the random variable X described in the previous section corresponded to the distance of a random raindrop from the centre of the circular pond.

Suppose one is interested in the *square* of the distance from the centre of the pond and how this derived value is distributed.

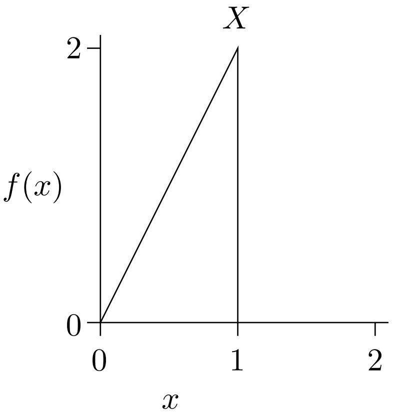
To investigate this, take the random variable X and apply to it the transformation function $y(x)$ specified as:

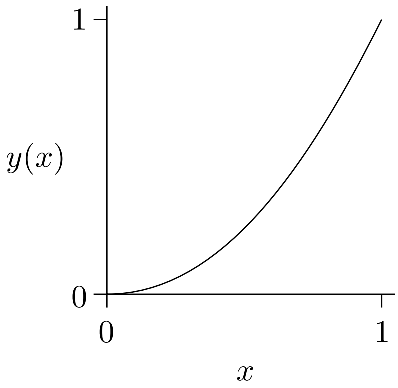
$$y(x) = x^2$$

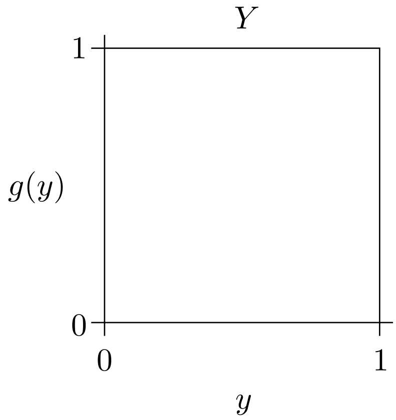
Note that the useful part of the range of x is 0 to 1 and, over this range, $y(x)$ increases monotonically from 0 to 1.

Let $Y = y(X)$ and let $g(y)$ be the probability density function associated with Y. What is the function $g(y)$?

The problem is illustrated by the following figures:







First, derive $x(y)$ the inverse of the function $y(x)$:

$$x(y) = \sqrt{y}$$

Accordingly:

$$f\big(x(y)\big) = 2\sqrt{y} \qquad \text{and} \qquad \frac{dx}{dy} = \frac{1}{2\sqrt{y}}$$

From (11.4):

$$g(y) = f\big(x(y)\big)\left|\frac{dx}{dy}\right| = 1$$

As illustrated in the figures, the function $y(x)$ transforms the triangular distribution $f(x)$ into the distribution $g(y)$ which is Uniform(0,1).

Transforming a Uniform Distribution into a Normal Distribution

It would be very useful if there were an easy way of transforming a uniform distribution into a normal distribution.

Suppose that X is a random variable whose distribution is Uniform(0,1) and Y is a random variable whose distribution is Normal(0,1). The associated probability density functions ($f(x)$ and $g(y)$ respectively) are:

$$f(x) = \begin{cases} 1, & \text{if } 0 \leqslant x < 1 \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad g(y) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2}$$

The goal is to determine a function $y(x)$ which will transform $f(x)$ into $g(y)$. Given that $f(x)$ is Uniform(0,1), relationship (11.5) above leads to the differential equation:

$$\frac{dx}{dy} = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2} \tag{11.6}$$

Unfortunately this differential equation is intractable.

Glossary

The following technical term has been introduced:

transformation function

Exercises — XI

1. Although (11.6) cannot be solved analytically, it succumbs to numerical methods. The required transformation function $y(x)$ is incorporated into Excel as the built-in function `NORMSINV`. Its use is illustrated in the Excel worksheet which is shown on the facing page. Prepare a worksheet like this one.

   The following steps are involved:

   (*a*) First set up a column of 99 values running from 0.01 to 0.99 in steps of 0.01. Only the first 51 of these values appear on the facing page (the worksheet runs to a second page which is not shown). Head this column x as shown.

   (*b*) Set up a second column headed $y(x)$. Each value is the result of applying the function `NORMSINV` to the corresponding value of x. Note that 0.0 and 1.0 are deliberately omitted as values of x because and $y(0) = -\infty$ and $y(1) = +\infty$. The range of the Uniform distribution is 0 to 1 and this maps into the range of the Normal distribution which is $-\infty$ to $+\infty$.

   (*c*) Use the chart wizard to set up the plot of the transformation function: $y(x)$ against x over the range of values 0.01 to 0.99. The chart only hints at how rapidly the function approaches $-\infty$ and $+\infty$ as x tends to zero or one.

   (*d*) Set up the column headed *Range*, whose 12 values run from -2.75 to 2.75. These values constitute the Bin Range required by the Histogram tool in Excel. . .

   (*e*) Check the Tools menu. If the Data Analysis command is not there, choose the Add-Ins command and, via that, pick up the Analysis ToolPak. Now choose the Data Analysis command and, via that, select the Histogram tool.

   Specify the range containing the 99 values under the heading $y(x)$ as the Input Range and specify the range containing the 12 values under the heading *Range* as the Bin Range. Select the Output Range option and, as the Output Range itself, specify the single cell two places to the right of the cell with *Range* in it. This will be the top left-hand cell of the table which the Histogram tool should then produce along with the lower chart.

   (*f*) Tidy up the chart and add comments to the worksheet. The overall result should have a neat appearance roughly as the worksheet opposite.

   Ideally the figures in the column headed *Frequency* should be half a row higher. The value 19 would then more clearly indicate that it is the number of values found *between* -0.25 and $+0.25$. The value 0 at the top is the number of values found *less than* -2.75 and the value 0 at the bottom is the number of values found *more than* $+2.75$ (hence the word More). The numbers against the x-axis of the chart are also rather unhappily placed.

   Despite these minor shortcomings the table and chart strongly suggest that a Uniform distribution has been transformed into a Normal distribution.

2. Replace the 99 values in the column headed x by `=RAND()`. The new values will be distributed Uniform(0,1) and the values in the $y(x)$ column will continue to be distributed Normal(0,1) but they will no longer be sorted. Ignore or delete the upper chart. Delete the table and the lower chart and invoke the Histogram tool again. The histogram which results will not be quite so convincing as its predecessor but it should not be very different.

3. Extend the two main columns so that instead of 99 pairs of values there are 1000. Delete the table and chart and invoke the Histogram tool (remember to extend the Input Range). Note that the results are again fairly convincing.

4. Rework the original worksheet (on the previous page) but replace the transformation function `NORMSINV` by $-\frac{1}{2}\ln x$ and invoke the Histogram tool once again. Check that the results are in reasonable accordance with Example II on page 11.7.

5. Given the probability density functions:

$$f(x) = \begin{cases} 1, & \text{if } 0 \leqslant x < 1 \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad g(y) = \begin{cases} \dfrac{y}{2}, & \text{if } 0 \leqslant y < 2 \\ 0, & \text{otherwise} \end{cases}$$

   determine the function $y(x)$ which will transform $f(x)$ into $g(y)$.

   Rework the worksheet again to illustrate the use of the derived transformation function to transform the uniform distribution whose probability density function is $f(x)$ into the triangular distribution whose probability density function is $g(y)$.

6. The triangular distribution obtained in the previous exercise is the same as that whose probability density function was given as $f(x)$ in Example I on page 11.4. By applying the transformation function used in Example I to the values in the second column, set up a third column whose values should be distributed in accordance with the triangular distribution obtained in Example I. Use the Histogram tool to demonstrate this.