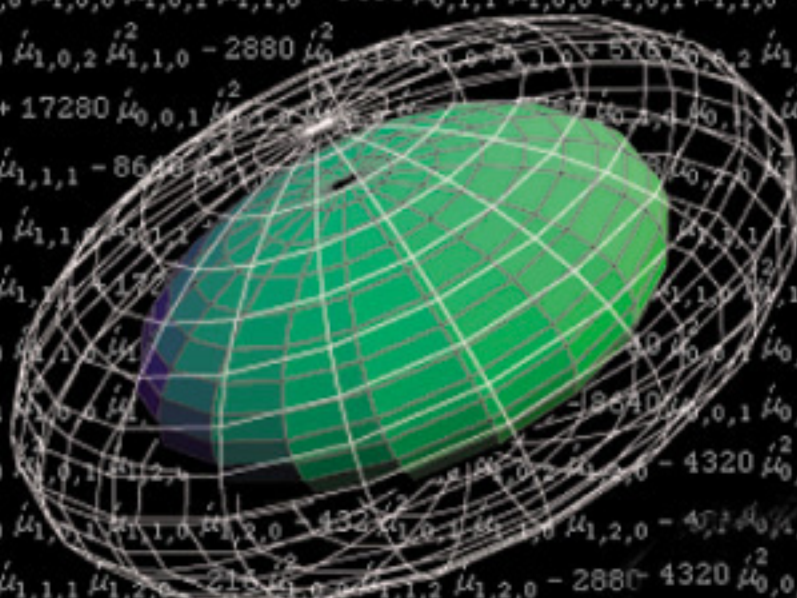
SPRINGER TEXTS IN STATISTICS

MATHEMATICAL STATISTICS *with Mathematica*®



COLIN ROSE
MURRAY D. SMITH

Mathematical Statistics with *Mathematica*

Chapter 4 – Distributions of Functions of Random Variables

Please reference this 2002 edition as:

Rose, C. and Smith, M.D. (2002)
Mathematical Statistics with Mathematica, Springer-Verlag, New York.

Latest edition

For the latest up-to-date edition, please visit: www.mathStatica.com

Chapter 4

Distributions of Functions of Random Variables

4.1 Introduction

This chapter is concerned with the following problem, which we state here in its simplest form:

> Let X be a random variable with density $f(x)$.
> What is the distribution of $Y = u(X)$, where $u(X)$ denotes some function of X?

This problem is of interest for several reasons. First, it is crucial to an understanding of statistical *distribution theory*: for instance, this chapter derives (from first principles) distributions such as the Lognormal, Pareto, Extreme Value, Rayleigh, Chi-squared, Student's t, Fisher's F, noncentral Chi-squared, noncentral F, Triangular and Laplace, amongst many others. Second, it is important in *sampling theory*: the chapter discusses ways to find the exact sampling distribution of statistics such as the sample sum, the sample mean, and the sample sum of squares. Third, it is of practical importance: for instance, a gold mine may have a profit function $u(x)$ that depends on the gold price X (a random variable). The firm is interested to know the distribution of its profits, given the distribution of X.

In statistics, there are two standard methods for solving these problems:

- The *Transformation Method*: this only applies to one-to-one transformations.
- The *MGF Method*: this is less restrictive, but can be more difficult to solve. It is based on the Uniqueness Theorem relating moment generating functions to densities.

§4.2 discusses the Transformation Method, while §4.3 covers the MGF Method. These two methodologies are then applied to some important examples in §4.4 (products and ratios of random variables) and §4.5 (sums and differences of random variables).

4.2 The Transformation Method

This section discusses the Transformation Method: §4.2 A discusses transformations of a single random variable, §4.2 B extends the analysis to the multivariate case, while §4.2 C considers transformations that are not strictly one-to-one, as well as manual methods.

4.2 A Univariate Cases

A *one-to-one transformation* implies that each value x is related to one (and only one) value $y = u(x)$, and that each value y is related to one (and only one) value $x = u^{-1}(y)$. Any univariate monotonic increasing or decreasing function yields a one-to-one transformation. Figure 1, for instance, shows two transformations.

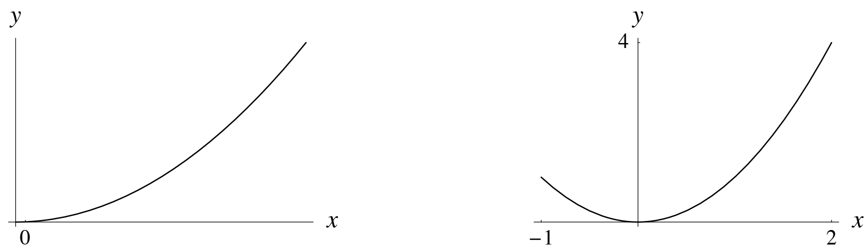


Fig. 1: (i) $y = x^2$, for $x \in \mathbb{R}_+$ (ii) $y = x^2$, for $x \in (-1, 2)$

Case (i): Even though $y = x^2$ has two solutions, namely:

```
Solve[y == x^2, x]
```

$$\{\{x \to -\sqrt{y}\}, \{x \to \sqrt{y}\}\}$$

only the latter solution is valid for the given domain ($x \in \mathbb{R}_+$). Therefore, *over the given domain*, the function is monotonically increasing, and thus case (i) is a one-to-one transformation.

Case (ii): Here, for some values of y, there exists more than one corresponding value of x; there are now two valid solutions, neither of which can be excluded. Thus, case (ii) is *not* a one-to-one transformation. Fortunately, a theorem exists to deal with such cases: see §4.2 C.

Theorem 1: Let X be a *continuous* random variable with pdf $f(x)$, and let $Y = u(X)$ define a one-to-one transformation between the values of X and Y. Then the pdf of Y, say $g(y)$, is

$$g(y) = f\left(u^{-1}(y)\right) |J| \tag{4.1}$$

where $x = u^{-1}(y)$ is the inverse function of $y = u(x)$, and $J = \frac{d\,u^{-1}(y)}{dy}$ denotes the Jacobian of the transformation; u^{-1} is assumed to be differentiable.

Proof: We will only sketch the proof.[1] To aid intuition, suppose $Y = u(X)$ defines a one-to-one *increasing* transformation between the values of X and Y. Then $P(Y \le y) = P(X \le x)$, or equivalently in terms of their respective cdf's, $G(y) = F(x)$. Then, by the chain rule of differentiation:

$$g(y) = \frac{dG(y)}{dy} = \frac{dF(x)}{dx}\frac{dx}{dy} = f(x)\frac{dx}{dy} \qquad \text{where } x = u^{-1}(y).$$

Remark: If X is a *discrete* random variable, then (4.1) becomes:

$$g(y) = f\left(u^{-1}(y)\right)$$

The **mathStatica** function, `Transform[`*eqn*, *f*`]` finds the density of $Y = u(X)$, where X has density $f(x)$, for both continuous and discrete random variables, while `TransformExtremum[`*eqn*, *f*`]` calculates the domain of Y, if it can do so. As per Theorem 1, `Transform` and `TransformExtremum` should only be used on transformations that are one-to-one. The `Transform` function is best illustrated by example …

⊕ ***Example 1:*** Derivation of the Cauchy Distribution

Let X have Uniform density $f(x) = \frac{1}{\pi}$, defined on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$:

$$\mathbf{f} = \frac{1}{\pi}; \qquad \mathbf{domain[f]} = \left\{x, -\frac{\pi}{2}, \frac{\pi}{2}\right\};$$

Then, the density of $Y = \tan(X)$ is derived as follows:

```
Transform[y == Tan[x], f]
```

$$\frac{1}{\pi + \pi\, y^2}$$

with domain of support:

```
TransformExtremum[y == Tan[x], f]
```

```
{y, -∞, ∞}
```

This is the pdf of a Cauchy distributed random variable. Note the double equal sign in the transformation equation: `y == Tan[x]`. If, by mistake, we enter `y = Tan[x]` with a single equal sign (or if `y` was previously given some value), we would need to `Clear[y]` before trying again. ■

⊕ ***Example 2:*** Standardising a $N(\mu, \sigma^2)$ Random Variable

Let $X \sim N(\mu, \sigma^2)$ with density $f(x)$:

```
f = e^(-(x-μ)^2/(2σ^2)) / (σ Sqrt[2π]);    domain[f] = {x, -∞, ∞} && {μ ∈ Reals, σ > 0};
```

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

Then, the density of $Z = \frac{X-\mu}{\sigma}$, denoted $g(z)$ is:

```
        g  = Transform[z == (x - μ)/σ, f]
domain[g]  = TransformExtremum[z == (x - μ)/σ, f]
```

$$\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}$$

{z, -∞, ∞}

That is, Z is a $N(0, 1)$ random variable. ■

⊕ ***Example 3:*** Derivation of the Lognormal Distribution

Let $X \sim N(\mu, \sigma^2)$ with density $f(x)$, as entered above in *Example 2*. Then, the density of $Y = e^X$, denoted $g(y)$, is:

```
        g  = Transform[y == e^x, f]
domain[g]  = TransformExtremum[y == e^x, f]
```

$$\frac{e^{-\frac{(\mu-\text{Log}[y])^2}{2\sigma^2}}}{\sqrt{2\pi}\, y\, \sigma}$$

{y, 0, ∞} && {μ ∈ Reals, σ > 0}

This is a Lognormal distribution, so named because $\log(Y)$ has a Normal distribution. Figure 2 plots the Lognormal pdf, when $\mu = 0$ and $\sigma = 1$.

```
PlotDensity[g /. {μ → 0, σ → 1}];
```

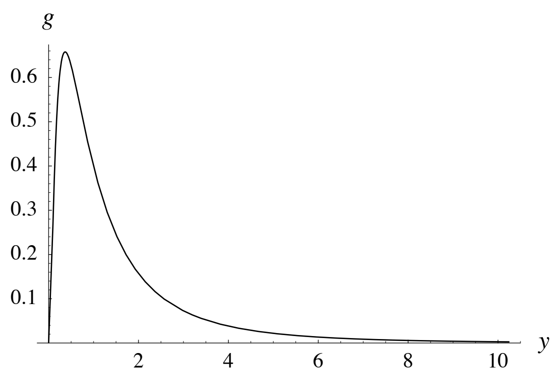


Fig. 2: Lognormal pdf

⊕ ***Example 4:*** Derivation of Uniform, Pareto, Extreme Value and Rayleigh Distributions

Let X have a standard Exponential distribution with density $f(x)$:

```
f = ⅇ^-x;      domain[f] = {x, 0, ∞};
```

We shall consider the following simple transformations:

$$\text{(i) } Y = e^{-X} \quad \text{(ii) } Y = e^{X} \quad \text{(iii) } Y = -\log(X) \quad \text{(iv) } Y = \sqrt{X}$$

(i) When $Y = e^{-X}$, we obtain the standard Uniform distribution:

```
        g  = Transform[y == ⅇ^-x, f]
domain[g] = TransformExtremum[y == ⅇ^-x, f]
```

1

{y, 0, 1}

(ii) When $Y = e^{X}$, we obtain a Pareto distribution:

```
        g  = Transform[y == ⅇ^x, f]
domain[g] = TransformExtremum[y == ⅇ^x, f]
```

$\frac{1}{y^2}$

{y, 1, ∞}

More generally, if $X \sim \text{Exponential}(\frac{1}{a})$, then $Y = b\, e^{X}$ $(b > 0)$ yields the Pareto density with pdf $a\, b^a\, y^{-(a+1)}$, defined for $y > b$.[2] This is often used in economics to model the distribution of income, and is named after the economist Vilfredo Pareto (1848–1923).

(iii) When $Y = -\log(X)$, we obtain the standard Extreme Value distribution:

```
        g  = Transform[y == -Log[x], f]
domain[g] = TransformExtremum[y == -Log[x], f]
```

$e^{-e^{-y}-y}$

{y, -∞, ∞}

(iv) When $Y = \sqrt{X}$, we obtain a Rayleigh distribution:

```
        g  = Transform[y == √x , f]
domain[g] = TransformExtremum[y == √x , f]
```

$2\, e^{-y^2}\, y$

$\{y, 0, \infty\}$

as given in the *Continuous* palette (simply replace σ with $\sqrt{1/2}$ to get the same result). More generally, if $X \sim \text{Exponential}(\lambda)$, then $Y = \sqrt{X} \sim \text{Rayleigh}(\sigma)$ with $\sigma = \sqrt{\lambda/2}$. This distribution is often used in engineering to model the life of electronic components. ■

⊕ ***Example 5:*** Transformations of the Uniform Distribution

Let $X \sim \text{Uniform}(\alpha, \beta)$ with density $f(x)$, where $0 < \alpha < \beta < \infty$:

$$f = \frac{1}{\beta - \alpha}; \quad \text{domain[f] = \{x, } \alpha, \beta\} \text{ \&\& } \{0 < \alpha < \beta\};$$

We seek the distributions of: (i) $Y = 1 + X^2$ and (ii) $Y = (1 + X)^{-1}$.

Solution: Let $g(y)$ denote the pdf of Y. Then the solution to (i) is:

```
        g  = Transform[y == 1 + x², f]
domain[g] = TransformExtremum[y == 1 + x², f]
```

$$\frac{1}{\sqrt{-1 + y}\ (-2\,\alpha + 2\,\beta)}$$

$\{y, 1 + \alpha^2, 1 + \beta^2\}$ && $\{0 < \alpha < \beta\}$

while the solution to the second part is:

```
        g  = Transform[y == (1 + x)^-1, f]
domain[g] = TransformExtremum[y == (1 + x)^-1, f]
```

$$\frac{1}{-y^2\,\alpha + y^2\,\beta}$$

$\left\{y, \frac{1}{1+\beta}, \frac{1}{1+\alpha}\right\}$ && $\{0 < \alpha < \beta\}$

Generally, transformations involving parameters pose no problem, provided we remember to attach the appropriate assumptions to the original `domain[f]` statement at the very start. ■

4.2 B Multivariate Cases

Thus far, we have considered the distribution of a transformation of a single random variable. This section extends the analysis to more than one random variable. The concepts discussed in the univariate case carry over to the multivariate case with the appropriate modifications.

Theorem 2: Let X_1 and X_2 be *continuous* random variables with joint pdf $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ define a one-to-one transformation between the values of (X_1, X_2) and (Y_1, Y_2). Then the joint pdf of Y_1 and Y_2 is

$$g(y_1, y_2) = f\left(u_1^{-1}(y_1, y_2),\ u_2^{-1}(y_1, y_2)\right) \ |J| \tag{4.2}$$

where $u_i^{-1}(y_1, y_2)$ is the inverse function of $Y_i = u_i(X_1, X_2)$, and

$$J = \det\begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{pmatrix}$$

is the Jacobian of the transformation, with $\frac{\partial x_i}{\partial y_j}$ denoting the partial derivative of $x_i = u_i^{-1}(y_1, y_2)$ with respect to y_j, and $\det(\cdot)$ denotes the determinant of the matrix. Transformations in higher dimensional systems follow in similar fashion.

Proof: The proof is analogous to Theorem 1; see Tjur (1980, §3.1) for more detail.

Remark: If the X_i are *discrete* random variables, (4.2) becomes:

$$g(y_1, y_2) = f\left(u_1^{-1}(y_1, y_2),\ u_2^{-1}(y_1, y_2)\right)$$

The **mathStatica** function, `Transform`, may also be used in multivariate settings. Of course, by Theorem 2, it should only be used to solve transformations that are one-to-one.

The transition from univariate to multivariate transformations raises two new issues:

(i) How many random variables?

The Transformation Method requires that there are as many 'new' variables Y_i as there are 'old' variables X_i. Suppose, for instance, that X_1, X_2 and X_3 have joint pdf $f(x_1, x_2, x_3)$, and that we seek the pdf of $Y_1 = u_1(X_1, X_2, X_3)$. This problem involves three steps. *First*, we must create two additional random variables, $Y_2 = u_2(X_1, X_2, X_3)$ and $Y_3 = u_3(X_1, X_2, X_3)$, and we must do so in such a way that there is one-to-one transformation from the values of (X_1, X_2, X_3) to (Y_1, Y_2, Y_3). This could, for example, be done by setting $Y_2 = X_2$, and $Y_3 = X_3$. *Second*, we can then find the joint pdf of (Y_1, Y_2, Y_3). *Third*, we can then derive the desired marginal pdf of Y_1 from the joint pdf of (Y_1, Y_2, Y_3) by integrating out Y_2 and Y_3. *Example 7* illustrates this procedure.

(ii) Non-rectangular domains

Let (X_1, X_2) have joint pdf $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ define a one-to-one transformation from the values of (X_1, X_2) to the values of (Y_1, Y_2), and let $g(y_1, y_2)$ denote the joint pdf of (Y_1, Y_2). Finally, let $\mathcal{A}$ denote the space where $f(x_1, x_2) > 0$, and let $\mathcal{B}$ denote the space where $g(y_1, y_2) > 0$; $\mathcal{A}$ and $\mathcal{B}$ are therefore the domains of support. Then, the transformation is said to map space $\mathcal{A}$ (in the x_1-x_2 plane) onto space $\mathcal{B}$ (in the y_1-y_2 plane). If the domain of a joint pdf does *not* depend on any of its constituent random variables, then we say the domain defines an *independent product space*. For instance, the domain $\mathcal{A} = \{(x_1, x_2): \frac{1}{2} < x_1 < 3,\ 1 < x_2 < 4\}$ is an independent product space, because the domain of X_1 does not depend on the domain of X_2, and vice versa. If plotted in x_1-x_2 space, this domain would appear rectangular, as the left panel in Fig. 3 illustrates.

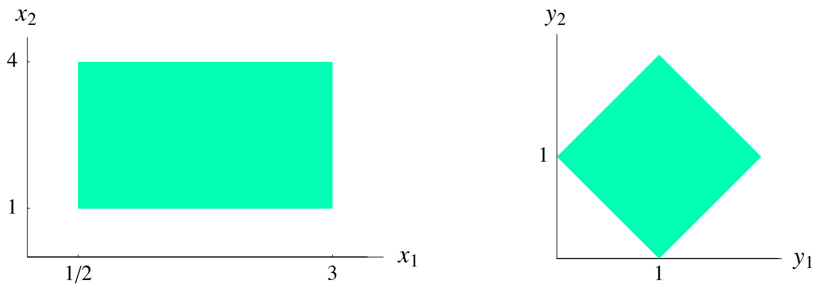


Fig. 3: Rectangular (left) and non-rectangular (right) domains

In this vein, we refer to domains as being either *rectangular* or *non-rectangular*. Even though space $\mathcal{A}$ is rectangular, it is important to realise that a multivariate transformation will often create dependence in space $\mathcal{B}$. To see this, consider the following example:

⊕ ***Example 6:*** A Non-Rectangular Domain

Let X_1 and X_2 be defined on the unit interval with joint pdf $f(x_1, x_2) = 1$:

```
f = 1;    domain[f] = {{x₁, 0, 1}, {x₂, 0, 1}};
```

Let $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. Then, we have:

```
eqn = {y₁ == x₁ + x₂, y₂ == x₁ - x₂};
```

Note the bracketing on the transformation equation — it takes the same form as *Mathematica*'s `Solve` function. Then the joint pdf of Y_1 and Y_2, denoted $g(y_1, y_2)$, is:

```
g = Transform[eqn, f]
```

$\frac{1}{2}$

The **mathStatica** function `DomainPlot[`*eqn*, *f*`]` illustrates set $\mathcal{B}$, denoting the space in the y_1-y_2 plane where $g(y_1, y_2) = \frac{1}{2}$.

```
DomainPlot[eqn, f];
```

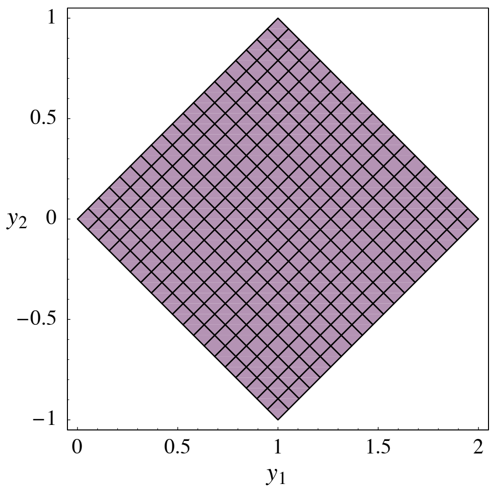


Fig. 4: Space in the y_1-y_2 plane where $g(y_1, y_2) = \frac{1}{2}$

The domain here is $\mathcal{B} = \{(y_1, y_2) : 0 < y_1 + y_2 < 2, \ -2 < y_2 - y_1 < 0\}$. This is clearly a non-rectangular domain, indicating that Y_1 and Y_2 are dependent.

Notes:

(i) In the multivariate case, `TransformExtremum` does not derive the domain itself; instead it calculates the extremities of the domain:

```
TransformExtremum[eqn, f]
```

```
{{y1, 0, 2}, {y2, -1, 1}}
```

This is sometimes helpful to verify ones working. However, as this example shows, extremities and domains are *not* always the same, and care must be taken not to confuse them.

(ii) For more information on `DomainPlot`, see the **mathStatica Help** file.

(iii) It is worth noting that even though Y_1 and Y_2 are dependent, they are uncorrelated:

```
Corr[{x1 + x2, x1 - x2}, f]
```

```
0
```

It follows that zero correlation does *not* imply independence. ■

⊕ ***Example 7:*** Product of Uniform Random Variables

Let $X_1 \sim$ Uniform(0, 1) be independent of $X_2 \sim$ Uniform(0, 1), and let $Y = X_1 X_2$. Find $P(Y \leq \frac{1}{4})$.

Solution: Due to independence, the joint pdf of X_1 and X_2, say $f(x_1, x_2)$, is just the pdf of X_1 multiplied by the pdf of X_2:

```
f = 1;     domain[f] = {{x₁, 0, 1}, {x₂, 0, 1}};
```

Take $Y = X_1 X_2$, and let $Z = X_2$, so that the number of 'new' variables is equal to the number of 'old' ones. Then, the transformation equation is:

```
eqn = {y == x₁ x₂,  z == x₂};
```

Let $g(y, z)$ denote the joint pdf of (Y, Z):

```
g = Transform[eqn, f]
```

$$\frac{1}{z}$$

Since X_1 and X_2 are $U(0, 1)$, and $Y = X_1 X_2$ and $Z = X_2$, it follows that $0 < y < z < 1$. To see this visually, evaluate `DomainPlot[eqn, f]`. We enter $0 < y < z < 1$ as follows:

```
domain[g] = {{y, 0, z}, {z, y, 1}} ;
```

Then the marginal pdf of $Y = X_1 X_2$ is:

```
h = Marginal[y, g]
```

```
-Log[y]
```

with domain of support:

```
domain[h] = {y, 0, 1};
```

Finally, we require $P(Y \leq \frac{1}{4})$. This is given by:

```
Prob[1/4, h]
```

$$\frac{1}{4}\ (1 + \text{Log}[4])$$

which is approximately 0.5966. It can be helpful, sometimes, to check that one's symbolic workings make sense by using an alternative methodology. For instance, we can use simulation to estimate $P(X_1 X_2 \leq \frac{1}{4})$. Here, then, are 10000 drawings of $Y = X_1 X_2$:

```
data = Table[ Random[] Random[], {10000}];
```

We now count how many copies of Y are smaller than (or equal to) $\frac{1}{4}$, and divide by 10000 to get our estimate of $P(Y \leq \frac{1}{4})$:

```
Count[data, y_ /; y ≤ 1/4]
--------------------------
         10000.
```

```
0.5952
```

which is close to the exact result derived above. ■

4.2 C Transformations That Are *Not* One-to-One; Manual Methods

In §4.2 A, we considered the transformation $Y = X^2$ defined on $x \in (-1, 2)$. This is *not* a one-to-one transformation, because for some values of Y there are two corresponding values of X. This section discusses how to undertake such transformations.

Theorem 3: Let X be a *continuous* random variable with pdf $f(x)$, and let $Y = u(X)$ define a transformation between the values of X and Y that is *not* one-to-one. Thus, if $\mathcal{A}$ denotes the space where $f(x) > 0$, and $\mathcal{B}$ denotes the space where $g(y) > 0$, then there exist points in $\mathcal{B}$ that correspond to more than one point in $\mathcal{A}$. However, if set $\mathcal{A}$ can be partitioned into k sets, $\mathcal{A}_1, \ldots, \mathcal{A}_k$, such that u defines a one-to-one transformation of each $\mathcal{A}_i$ onto $\mathcal{B}_i$ (the image of $\mathcal{A}_i$ under u), then the pdf of Y is

$$g(y) = \sum_{i=1}^{k} \delta_i(y)\, f\big(u_i^{-1}(y)\big)\, \big|J_i\big| \qquad \text{for } i = 1, \ldots, k \tag{4.3}$$

where $\delta_i(y) = 1$ if $y_i \in \mathcal{B}_i$ and 0 otherwise, $x = u_i^{-1}(y)$ is the inverse function of $Y = u(X)$ in partition i, and $J_i = \frac{d u_i^{-1}(y)}{dy}$ denotes the Jacobian of the transformation in partition i.[3]

All this really means is that, for each region i, we simply work as we did before with Theorem 1; we then add up all the parts $i = 1, \ldots, k$.

⊕ ***Example 8:*** A Transformation That Is *Not* One-to-One

Let X have pdf $f(x) = \frac{e^x}{e^2 - e^{-1}}$ defined on $x \in (-1, 2)$, and let $Y = X^2$. We seek the pdf of Y. We have:

$$f = \frac{e^x}{e^2 - e^{-1}}; \qquad \texttt{domain[f] = \{x, -1, 2\};} \qquad \texttt{eqn = \{y == x}^2\texttt{\};}$$

Solution: The transformation from X to Y is not one-to-one over the given domain. We can, however, partition the domain into two sets of points that are both one-to-one. We do this as follows:

```
f₁ = f;  domain[f₁] = {x, -1, 0};
f₂ = f;  domain[f₂] = {x,  0, 2};
```

Let $g_1(y)$ denote the density of Y corresponding to when $x \leq 0$, and similarly, let $g_2(y)$ denote the density of Y corresponding to $x > 0$:

```
{g₁ = Transform[eqn, f₁], TransformExtremum[eqn, f₁]}
{g₂ = Transform[eqn, f₂], TransformExtremum[eqn, f₂]}
```

$$\left\{\frac{e^{1-\sqrt{y}}}{(-2+2\,e^3)\sqrt{y}},\ \{y,\ 0,\ 1\}\right\}$$

$$\left\{\frac{e^{1+\sqrt{y}}}{(-2+2\,e^3)\sqrt{y}},\ \{y,\ 0,\ 4\}\right\}$$

By (4.3), it follows that

$$g(y) = \begin{cases} g_1 + g_2 & 0 < y \leq 1 \\ g_2 & 1 < y < 4 \end{cases}$$

which we enter, using **mathStatica**, as:

```
g = If[y ≤ 1, g₁ + g₂, g₂];  domain[g] = {y, 0, 4};
```

Figure 5 plots the pdf.

```
PlotDensity[g, PlotRange → {0, .5}];
```

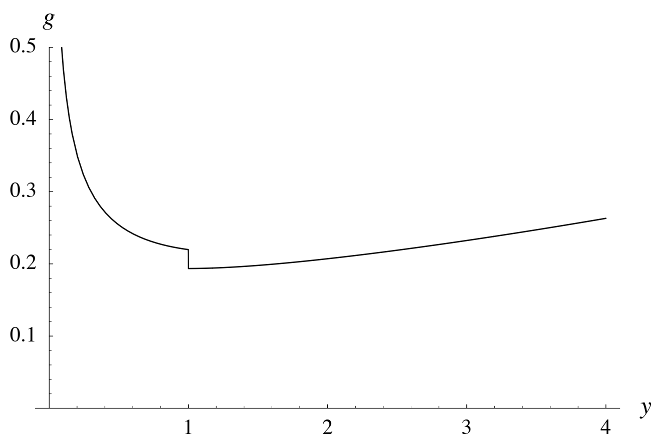


Fig. 5: The pdf of $Y = X^2$, with discontinuity at $y = 1$

Despite the discontinuity of the pdf at $y = 1$, **mathStatica** functions such as `Prob` and `Expect` will still work perfectly well. For instance, here is the cdf $P(Y \leq y)$:

```
cdf = Prob[y, g]
```

$$\text{If}\left[y \leq 1,\ \frac{2\,e\,\text{Sinh}\left[\sqrt{y}\right]}{-1+e^3},\ \frac{-1+e^{1+\sqrt{y}}}{-1+e^3}\right]$$

This can be easily illustrated with `Plot[cdf,{y,0,4}]`. ■

⊕ ***Example 9:*** The Square of a Normal Random Variable: The Chi-squared Distribution

Let $X \sim N(0, 1)$ with density $f(x)$. We seek the distribution of $Y = X^2$. Thus, we have:

$$\mathbf{f} = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}};\qquad \textbf{domain[f] = \{x, -}\infty\textbf{, }\infty\textbf{\};}\qquad \textbf{eqn = \{y == }x^2\textbf{\};}$$

Solution: The transformation equation here is *not* one-to-one over the given domain. By Theorem 3, we can, however, partition the domain into two disjoint sets of points that are both one-to-one:

```
f1 = f;   domain[f1] = {x, -∞, 0};
f2 = f;   domain[f2] = {x,  0, ∞};
```

Let $g_1(y)$ denote the density of Y corresponding to when $x \leq 0$, and similarly, let $g_2(y)$ denote the density of Y corresponding to when $x > 0$:

```
{g1 = Transform[eqn, f1], TransformExtremum[eqn, f1]}
{g2 = Transform[eqn, f2], TransformExtremum[eqn, f2]}
```

$$\left\{\frac{e^{-y/2}}{2\sqrt{2\pi}\,\sqrt{y}},\ \{y,\ 0,\ \infty\}\right\}$$

$$\left\{\frac{e^{-y/2}}{2\sqrt{2\pi}\,\sqrt{y}},\ \{y,\ 0,\ \infty\}\right\}$$

By Theorem 3, it follows that

$$g(y) = \begin{cases} g_1 + g_2 & 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

where $g_1 + g_2$ is:

```
g1 + g2
```

$$\frac{e^{-y/2}}{\sqrt{2\pi}\,\sqrt{y}}$$

This is the pdf of a Chi-squared random variable with 1 degree of freedom. ■

○ ***Manual Methods***

In all the examples above, we have always posed the transformation problem as:

Q. Let X be a random variable with pdf $f(x)$. What is the pdf of $Y = e^X$?
A. `Transform[y == ⅇ^x, f]`

But what if the same problem is posed as follows?

Q. Let X be a random variable with pdf $f(x)$. What is the pdf of Y, given $X = \log(Y)$?
A. `Transform[x == Log[y], f]` will fail, as this syntax is not supported.

We are now left with two possibilities:

(i) We could simply invert the transformation equation manually in *Mathematica* with `Solve[x == Log[y], f]`, and then derive the solution automatically with `Transform[y == ⅇ^x, f]`. Unfortunately, *Mathematica* may not always be able to neatly invert the transformation equation into the desired form, and we are then stuck.

(ii) Alternatively, we could adopt a manual approach by implementing either Theorem 1 (§4.2 A) or Theorem 2 (§4.2 B) ourselves in *Mathematica*. In a univariate setting, the basic approach would be to define:

```
g = (f /. x → Log[y]) * Jacob[x /. x → Log[y], y]
```

where the **mathStatica** function `Jacob` calculates the Jacobian of the transformation in absolute value. A multivariate example of a manual step-by-step transformation is given in Chapter 6 (see *Example 20*, §6.4 A).

4.3 The MGF Method

The moment generating function (mgf) method is based on the Uniqueness Theorem (§2.4 D) which states that there is a one-to-one correspondence between the mgf and the pdf of a random variable (if the mgf exists). Thus, if two mgf's are the same, then they must share the same density. As before, let X have density $f(x)$, and consider the transformation to $Y = u(X)$. We seek the pdf of Y, say $g(y)$. Two steps are involved:

Step 1: Find the mgf of Y.

Step 2: Hence, find the pdf of Y. This is normally done by matching the functional form of the mgf of Y with well-known moment generating functions. One usually does this armed with a textbook that lists the mgf's for well-known distributions, unless one has a fine memory for such things. If we can find a match, then the pdf is identified by the Uniqueness Theorem. Unfortunately, this matching process is often neither easy nor obvious. Moreover, if the pdf of Y is not well-known, then matching may not be possible. The mgf method is particularly well-suited to deriving the distribution of sample sums and sample means. This is discussed in §4.5 B, which provides further examples.

⊕ ***Example 10:*** The Square of a Normal Random Variable (again)

Let random variable $X \sim N(0, 1)$ with pdf $f(x)$:

```
f = e^(-x^2/2) / Sqrt[2 π];   domain[f] = {x, -∞, ∞};
```

We seek the distribution of $Y = X^2$.

Solution: The mgf of $Y = X^2$ is given by $E\left[e^{t X^2}\right]$:

```
mgf_Y = Expect[e^(t x^2), f]
```

— This further assumes that: $\left\{t < \frac{1}{2}\right\}$

$$\frac{1}{\sqrt{1 - 2\,t}}$$

By referring to a listing of mgf's, we see that this output is identical to the mgf of a Chi-squared random variable with 1 degree of freedom, confirming what was found in *Example 9*. Hence, if $X \sim N(0, 1)$, then X^2 is Chi-squared with 1 degree of freedom.

Using Characteristic Functions

The Uniqueness Theorem applies to both the moment generating function and the characteristic function (cf). As such, instead of deriving the mgf of Y, we could just as well have derived the characteristic function. Indeed, using the cf has two advantages. First, for many densities, the mgf does not exist, whereas the cf does. Second, once we have the cf, we can (in theory) derive the pdf that is associated with it by means of the Inversion Theorem (§2.4 D), rather than trying to match it with a known cf in a textbook appendix. This is particularly important if the derived cf is not of a standard (or common) form.

In this vein, we now obtain the pdf of Y directly by the Inversion Theorem. To start, we need the cf. Since we already know the mgf (derived above), we can easily derive the cf by simply replacing the argument t with $i\,t$, as follows:

```
cf = mgf_Y /. t → ⅈ t
```

$$\frac{1}{\sqrt{1 - 2\,i\,t}}$$

and then apply the Inversion Theorem (as per §2.4 D) to yield the pdf:

```
pdf = InverseFourierTransform[cf, t, y,
                FourierParameters → {1, 1}]
```

$$\frac{(1 + \mathrm{Sign}[y])\,\left(\mathrm{Cosh}\left[\frac{y}{2}\right] - \mathrm{Sinh}\left[\frac{y}{2}\right]\right)}{2\,\sqrt{2\,\pi}\,\left(y^2\right)^{1/4}}$$

which simplifies further if we note that Y is always positive:

```
FullSimplify[pdf, y > 0]
```

$$\frac{e^{-y/2}}{\sqrt{2\pi}\sqrt{y}}$$

which is the pdf we obtained in *Example 9*. Although inverting the cf is much more attractive than matching mgf's with textbook appendices, the inversion process is computationally difficult (even with *Mathematica*) and success is not that common in practice. ■

⊕ ***Example 11:*** Product of Two Normals

Let X_1 and X_2 be independent $N(0, 1)$ random variables. We wish to find the density of the product $Y = X_1 X_2$ using the mgf/cf method.

Solution: The joint pdf $f(x_1, x_2)$ is:

$$f = \frac{e^{-\frac{x_1^2}{2}}}{\sqrt{2\pi}} \frac{e^{-\frac{x_2^2}{2}}}{\sqrt{2\pi}};$$

```
domain[f] = {{x1, -∞, ∞}, {x2, -∞, ∞}};
```

The cf of Y is given by $E\left[e^{itY}\right] = E\left[e^{itX_1 X_2}\right]$:

```
cf = Expect[e^(i t x1 x2), f]
```

– This further assumes that: $\{t^2 > -1\}$

$$\frac{1}{\sqrt{1+t^2}}$$

Inverting the cf yields the pdf of Y:

```
pdf = InverseFourierTransform[cf, t, y,
            FourierParameters → {1, 1}]
```

$$\frac{\text{BesselK}[0,\ y\,\text{Sign}[y]]}{\pi}$$

where `BesselK` denotes the modified Bessel function of the second kind. Figure 6 contrasts the pdf of Y with that of the Normal pdf.

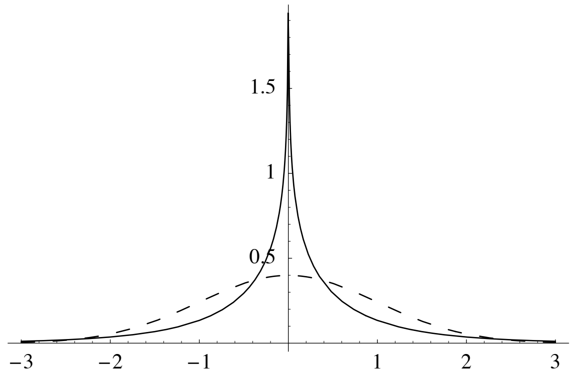


Fig. 6: The pdf of the product of two Normals (——) compared to a Normal pdf (– – –)

4.4 Products and Ratios of Random Variables

This section discusses random variables that are formed as products or ratios of other random variables.

⊕ ***Example 12:*** Product of Two Normals (again)

Let X_1 and X_2 be two independent standard Normal random variables. In *Example 11*, we found the pdf of the product $X_1\, X_2$ using the MGF Method. We now do so using the Transformation Method.

Solution: Let $f(x_1, x_2)$ denote the joint pdf of X_1 and X_2. Due to independence, $f(x_1, x_2)$ is just the pdf of X_1 multiplied by the pdf of X_2:

$$\mathbf{f} = \frac{e^{-\frac{x_1^2}{2}}}{\sqrt{2\pi}}\,\frac{e^{-\frac{x_2^2}{2}}}{\sqrt{2\pi}};\quad \mathbf{domain[f]} = \{\{x_1, -\infty, \infty\}, \{x_2, -\infty, \infty\}\};$$

Let $Y_1 = X_1\, X_2$ and $Y_2 = X_2$. Then, the joint pdf of (Y_1, Y_2), say $g(y_1, y_2)$, is:

$$\begin{aligned} \mathbf{g} &= \mathbf{Transform}[\{y_1 == x_1\, x_2,\ y_2 == x_2\},\ f]; \\ \mathbf{domain[g]} &= \{\{y_1, -\infty, \infty\},\ \{y_2, -\infty, \infty\}\}; \end{aligned}$$

In the interest of brevity, we have suppressed the output of `g` here by putting a semi-colon at the end of each line of the input. Nevertheless, one should always inspect the solution for `g` by removing the semi-colon, before proceeding further. Given $g(y_1, y_2)$, the marginal pdf of Y_1 is:

$$\mathbf{Marginal}[y_1,\ g]$$

$$\frac{\text{BesselK}[0,\ \text{Abs}[y_1]]}{\pi}$$

as per *Example 11*. ■

⊕ ***Example 13:*** Ratio of Two Normals: The Cauchy Distribution

Let X_1 and X_2 be two independent standard Normal random variables. We wish to find the pdf of the ratio X_1 / X_2.

Solution: The joint pdf $f(x_1, x_2)$ was entered in *Example 12*. Let $g(y_1, y_2)$ denote the joint pdf of $Y_1 = X_1 / X_2$ and $Y_2 = X_2$. Then:

$$\texttt{g} \;=\; \texttt{Transform}\left[\left\{y_1 == \frac{x_1}{x_2},\; y_2 == x_2\right\},\; f\right];$$
$$\texttt{domain[g]} \;=\; \{\{y_1, -\infty, \infty\},\; \{y_2, -\infty, \infty\}\};$$

Again, one should inspect the solution to g by removing the semi-colons. The pdf of Y_1 is:

```
Marginal[y1, g]
```

$$\frac{1}{\pi + \pi\, y_1^2}$$

where Y_1 has domain of support $(-\infty, \infty)$. That is, the ratio of two independent $N(0, 1)$ random variables has a Cauchy distribution. ■

⊕ ***Example 14:*** Derivation of Student's *t* Distribution

Let $X \sim N(0, 1)$ be independent of $Y \sim$ Chi-squared(n). We seek the density of the (scaled) ratio $T = \frac{X}{\sqrt{Y/n}}$.

Solution: Due to independence, the joint pdf of (X, Y), say $f(x, y)$, is the pdf of X multiplied by the pdf of Y:

$$\texttt{f} \;=\; \left(\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\right) * \left(\frac{y^{\frac{n}{2}-1}\, e^{-\frac{y}{2}}}{2^{n/2}\, \Gamma[\frac{n}{2}]}\right);$$
$$\texttt{domain[f]} \;=\; \{\{x, -\infty, \infty\},\; \{y, 0, \infty\}\}\; \&\&\; \{n > 0\};$$

Let $T = \frac{X}{\sqrt{Y/n}}$ and $Z = Y$. Then, the joint pdf of (T, Z), say $g(t, z)$, is obtained with:

$$\texttt{g} \;=\; \texttt{Transform}\left[\left\{t == \frac{x}{\sqrt{y/n}},\; z == y\right\},\; f\right];$$
$$\texttt{domain[g]} \;=\; \{\{t, -\infty, \infty\},\; \{z, 0, \infty\}\}\; \&\&\; \{n > 0\};$$

Then, the pdf of T is:

```
Marginal[t, g]
```

$$\frac{n^{n/2}\; (n + t^2)^{\frac{1}{2}(-1-n)}\; \Gamma[\frac{1+n}{2}]}{\sqrt{\pi}\; \Gamma[\frac{n}{2}]}$$

where T has domain of support $(-\infty, \infty)$. This is the pdf of a random variable distributed according to Student's *t* distribution with n degrees of freedom. ■

⊕ ***Example 15:*** Derivation of Fisher's F Distribution

Let $X_1 \sim \chi_a^2$ be independent of $X_2 \sim \chi_b^2$, where χ_a^2 and χ_b^2 are Chi-squared distributions with degrees of freedom a and b, respectively. We seek the distribution of the (scaled) ratio $R = \frac{X_1/a}{X_2/b}$.

Solution: Due to independence, the joint pdf of (X_1, X_2), say $f(x_1, x_2)$, is just the pdf of X_1 multiplied by the pdf of X_2:

$$\mathbf{f} = \left(\frac{x_1^{\frac{a}{2}-1}\, e^{-\frac{x_1}{2}}}{2^{\frac{a}{2}}\, \Gamma[\frac{a}{2}]}\right) * \left(\frac{x_2^{\frac{b}{2}-1}\, e^{-\frac{x_2}{2}}}{2^{\frac{b}{2}}\, \Gamma[\frac{b}{2}]}\right);$$

```
domain[f] = {{x₁, 0, ∞}, {x₂, 0, ∞}} && {a > 0, b > 0};
```

Let $Z = X_2$. Then, the joint pdf of (R, Z), say $g(r, z)$, is obtained with:

$$\mathbf{g} = \mathbf{Transform}\left[\left\{r == \frac{x_1 / a}{x_2 / b},\ z == x_2\right\},\ f\right];$$

```
domain[g] = {{r, 0, ∞}, {z, 0, ∞}} && {a > 0, b > 0};
```

Then, the pdf of random variable R is:

```
Marginal[r, g]
```

$$\frac{\left(\frac{a\,r}{b}\right)^{a/2} \left(1+\frac{a\,r}{b}\right)^{\frac{1}{2}(-a-b)} \Gamma[\frac{a+b}{2}]}{r\, \Gamma[\frac{a}{2}]\, \Gamma[\frac{b}{2}]}$$

with domain of support $(0, \infty)$. This is the pdf of a random variable with Fisher's F distribution, with parameters a and b denoting the numerator and denominator degrees of freedom, respectively. ■

⊕ ***Example 16:*** Derivation of Noncentral F Distribution

Let $X_1 \sim \chi_a^2(\lambda)$ be independent of $X_2 \sim \chi_b^2$, where $\chi_a^2(\lambda)$ denotes a noncentral Chi-squared distribution with noncentrality parameter λ. We seek the distribution of the (scaled) ratio $R = \frac{X_1/a}{X_2/b}$.

Solution: Let $f(x_1, x_2)$ denote the joint pdf of X_1 and X_2. Due to independence, $f(x_1, x_2)$ is just the pdf of X_1 multiplied by the pdf of X_2. As usual, the *Continuous* palette can be used to help enter the densities:

$$\mathbf{f} = \left(2^{-a/2}\, e^{-(x_1+\lambda)/2}\, x_1^{(a-2)/2} * \mathbf{Hypergeometric0F1Regularized}\left[\frac{a}{2}, \frac{x_1\,\lambda}{4}\right]\right) \left(\frac{x_2^{\frac{b}{2}-1}\, e^{-\frac{x_2}{2}}}{2^{\frac{b}{2}}\, \Gamma[\frac{b}{2}]}\right);$$

```
domain[f] = {{x₁, 0, ∞}, {x₂, 0, ∞}} && {a > 0, b > 0, λ > 0};
```

With $Z = X_2$, the joint pdf of (R, Z), say $g(r, z)$, is obtained with:

```
g  =  Transform[{r == (x1 / a)/(x2 / b),  z == x2},  f];
domain[g] = {{r, 0, ∞}, {z, 0, ∞}} && {a > 0, b > 0, λ > 0};
```

Then, the pdf of random variable R is:

```
Marginal[r, g]
```

$$\frac{1}{r\,\Gamma[\frac{b}{2}]}\left(e^{-\lambda/2}\left(\frac{a\,r}{b}\right)^{a/2}\left(1+\frac{a\,r}{b}\right)^{\frac{1}{2}(-a-b)}\Gamma\left[\frac{a+b}{2}\right]\right.$$
$$\left.\text{Hypergeometric1F1Regularized}\left[\frac{a+b}{2},\ \frac{a}{2},\ \frac{a\,r\,\lambda}{2\,b+2\,a\,r}\right]\right)$$

with domain of support $(0, \infty)$. This is the pdf of a random variable with a noncentral F distribution with noncentrality parameter λ, and degrees of freedom a and b. ■

4.5 Sums and Differences of Random Variables

This section discusses random variables that are formed as sums or differences of other random variables. §4.5 A applies the Transformation Method, while §4.5 B applies the MGF Method which is particularly well-suited to dealing with sample sums and sample means.

4.5 A Applying the Transformation Method

⊕ ***Example 17:*** Sum of Two Exponential Random Variables

Let X_1 and X_2 be independent random variables, each distributed Exponentially with parameter λ. We wish to find the density of $X_1 + X_2$.

Solution: Let $f(x_1, x_2)$ denote the joint pdf of (X_1, X_2):

$$f = \frac{e^{-\frac{x_1}{\lambda}}}{\lambda} * \frac{e^{-\frac{x_2}{\lambda}}}{\lambda};\quad \text{domain}[f] = \{\{x_1,\ 0,\ \infty\},\ \{x_2,\ 0,\ \infty\}\};$$

Let $Y = X_1 + X_2$ and $Z = X_2$. Since X_1 and X_2 are positive, it follows that $0 < z < y < \infty$. Then the joint pdf of (Y, Z), say $g(y, z)$, is obtained with:

```
g  =  Transform[{y == x1 + x2,  z == x2},  f];
domain[g]  =  {{y, z, ∞}, {z, 0, y}};
```

Then, the pdf of $Y = X_1 + X_2$ is:

```
Marginal[y, g]
```

$$\frac{e^{-\frac{y}{\lambda}}\, y}{\lambda^2}$$

with domain of support $(0, \infty)$, which is the pdf of a random variable with a Gamma distribution with shape parameter $a = 2$, and scale parameter $b = \lambda$. This is easy to verify using **mathStatica**'s *Continuous* palette. ■

⊕ ***Example 18:*** Sum of Poisson Random Variables

Let $X_1 \sim \text{Poisson}(\lambda_1)$ be independent of $X_2 \sim \text{Poisson}(\lambda_2)$. We seek the distribution of the sum $X_1 + X_2$.

Solution: Let $f(x_1, x_2)$ denote the joint pmf of (X_1, X_2):

$$\mathbf{f} = \frac{e^{-\lambda_1}\, \lambda_1^{x_1}}{x_1\,!}\ \frac{e^{-\lambda_2}\, \lambda_2^{x_2}}{x_2\,!};$$

$$\mathbf{domain[f] = \{\{x_1, 0, \infty\}, \{x_2, 0, \infty\}\}\ \&\&\ \{Discrete\};}$$

Let $Y = X_1 + X_2$ and $Z = X_2$. Then the joint pmf of (Y, Z), say $g(y, z)$, is:

```
g = Transform[{y == x₁ + x₂, z == x₂}, f]
```

$$\frac{e^{-\lambda_1-\lambda_2}\, \lambda_1^{y-z}\, \lambda_2^{z}}{(y-z)\,!\, z\,!}$$

where $0 \le z \le y < \infty$. We seek the pmf of Y, and so sum out the values of Z:

$$\mathbf{sol} = \sum_{z=0}^{y} \mathbf{Evaluate[g]}$$

$$\frac{e^{-\lambda_1-\lambda_2}\, \lambda_1^{y}\, \left(\frac{\lambda_1+\lambda_2}{\lambda_1}\right)^{y}}{\Gamma[1+y]}$$

which simplifies further:

```
FullSimplify[sol, y ∈ Integers]
```

$$\frac{e^{-\lambda_1-\lambda_2}\, (\lambda_1+\lambda_2)^{y}}{\Gamma[1+y]}$$

This is the pmf of a $\text{Poisson}(\lambda_1 + \lambda_2)$ random variable. Thus, the sum of independent Poisson variables is itself Poisson distributed. This result is particularly important in the following scenario: consider the sample sum comprised of n independent $\text{Poisson}(\lambda)$ variables. Then, $\sum_{i=1}^{n} X_i \sim \text{Poisson}(n\,\lambda)$. ■

⊕ ***Example 19:*** Sum of Two Uniform Random Variables: A Triangular Distribution

Let $X_1 \sim$ Uniform(0, 1) be independent of $X_2 \sim$ Uniform(0, 1). We seek the density of $Y = X_1 + X_2$.

Solution: Let $f(x_1, x_2)$ denote the joint pdf of (X_1, X_2):

```
f = 1;    domain[f] = {{x1, 0, 1}, {x2, 0, 1}};
```

Let $Y = X_1 + X_2$ and $Z = X_2$. Then the joint pdf of (Y, Z), say $g(y, z)$, is:

```
eqn = {y == x1 + x2, z == x2};    g = Transform[eqn, f]
1
```

Deriving the domain of this joint pdf is a bit more tricky, but can be assisted by using `DomainPlot`, which again plots the space in the y-z plane where $g(y, z) > 0$:

```
DomainPlot[eqn, f];
```

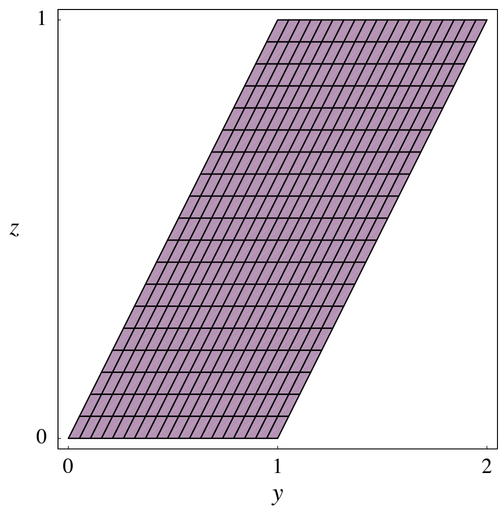


Fig. 7: The space in the y-z plane where $g(y, z) > 0$

We see that the domain (the shaded region) can be defined as follows:

When $y < 1$: $0 < z < y < 1$
When $y > 1$: $1 < y < 1 + z < 2$, or equivalently, $0 < y - 1 < z < 1$

The density of Y, say $h(y)$, is then obtained by integrating out Z in each part of the domain. This is easiest to do manually here:

$$h = \text{If}\left[y < 1,\ \text{Evaluate}\left[\int_0^y g\, dz\right],\ \text{Evaluate}\left[\int_{y-1}^1 g\, dz\right]\right]$$

```
If[y < 1, y, 2 - y]
```

with domain of support:

```
domain[h] = {y, 0, 2};
```

Figure 8 plots the pdf of Y.

```
PlotDensity[h];
```

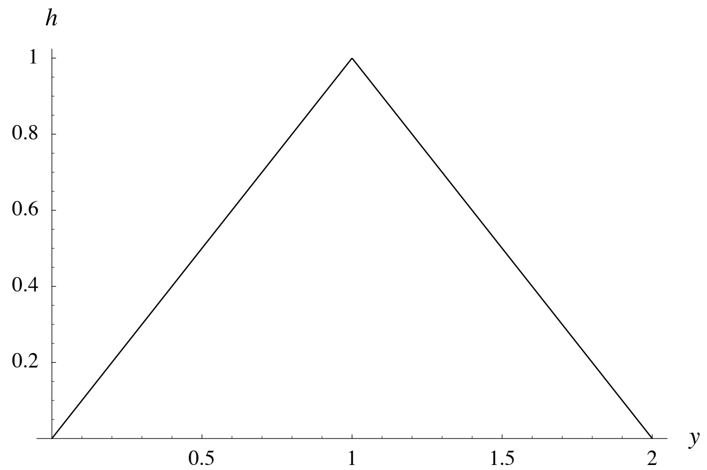


Fig. 8: Triangular pdf

This is known as a Triangular distribution. More generally, if $X_1, \ldots, X_n$ are independent Uniform(0,1) random variables, the distribution of $S_n = \sum_{i=1}^{n} X_i$ is known as the Irwin–Hall distribution (see *Example 18* of Chapter 2). By contrast, the distribution of S_n / n is known as Bates's distribution (*cf. Example 6* of Chapter 8). ■

⊕ ***Example 20:*** Difference of Exponential Random Variables: The Laplace Distribution

Let X_1 and X_2 be independent random variables, each distributed Exponentially with parameter $\lambda = 1$. We seek the density of $Y = X_1 - X_2$.

Solution: Let $f(x_1, x_2)$ denote the joint pdf of X_1 and X_2. Due to independence:

```
f = ⅇ^-x₁ * ⅇ^-x₂;   domain[f] = {{x₁, 0, ∞}, {x₂, 0, ∞}};
```

Let $Z = X_2$. Then the joint pdf of (Y, Z), say $g(y, z)$, is:

```
eqn = {y == x₁ - x₂, z == x₂};        g = Transform[eqn, f]
```

e^{-y-2z}

Deriving the domain of support of Y and Z is a bit more tricky. To make things clearer, we again use `DomainPlot` to plot the space in the y-z plane where $g(y, z) > 0$. Because x_1

and x_2 are unbounded above, we need to manually specify the plot bounds; we use `{x₁, 0, 100}, {x₂, 0, 100}` here:

```
DomainPlot[eqn, f, {x₁, 0, 100}, {x₂, 0, 100},
  PlotRange → {{-2, 2}, {-1, 3}}];
```

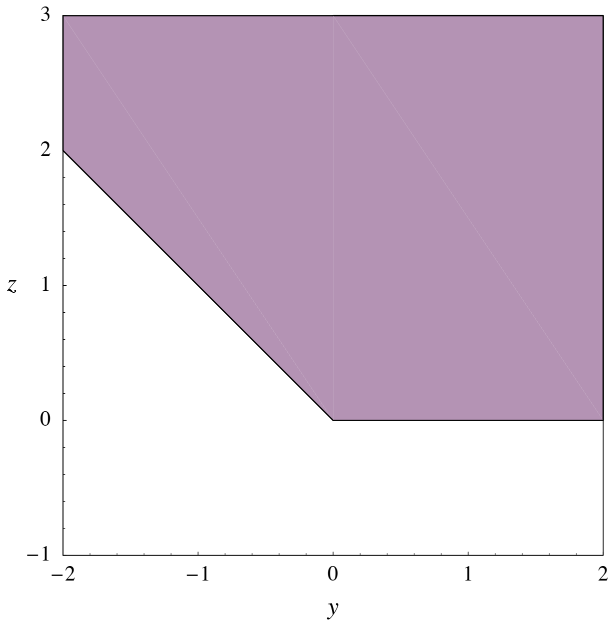


Fig. 9: The domain of support of Y and Z

This suggests that the domain (the shaded region in Fig. 9) can be defined as follows:

When $y < 0$: $0 < -y \le z < \infty$
When $y > 0$: $\{0 < y < \infty,\ 0 < z < \infty\}$

The density of Y, say $h(y)$, is then obtained by integrating out Z in each part of the domain. This is done manually here:

$$\mathbf{h = If}\left[\mathbf{y < 0,\ Evaluate}\left[\int_{-y}^{\infty} \mathbf{g}\, d\mathbf{z}\right],\ \mathbf{Evaluate}\left[\int_{0}^{\infty} \mathbf{g}\, d\mathbf{z}\right]\right]$$

$$\text{If}\left[y < 0,\ \frac{e^{y}}{2},\ \frac{e^{-y}}{2}\right]$$

with domain of support:

```
domain[h] = {y, -∞, ∞};
```

This is often expressed in texts as $h(y) = \frac{1}{2} e^{-|y|}$, for $y \in \mathbb{R}$. This is the pdf of a random variable with a standard Laplace distribution (also known as the Double Exponential distribution). ■

4.5 B Applying the MGF Method

The MGF Method is especially well-suited to finding the distribution of the sum of independent and identical random variables. Let $(X_1, \ldots, X_n)$ denote a random sample of size n drawn from a random variable X whose mgf is $M_X(t)$. Further, let:

$$\begin{aligned} s_1 &= \sum_{i=1}^{n} X_i \qquad \text{(sample sum)} \\ s_2 &= \sum_{i=1}^{n} X_i^2 \qquad \text{(sample sum of squares)} \end{aligned} \tag{4.4}$$

Then, the following results are a special case of the MGF Theorem of Chapter 2:

$$\begin{aligned} &\text{mgf of } s_1: && M_{s_1}(t) = \prod_{i=1}^{n} M_X(t) = \{M_X(t)\}^n = \left(E\left[e^{tX}\right]\right)^n \\ &\text{mgf of } \overline{X} = \tfrac{s_1}{n}: && M_{\overline{X}}(t) = M_{s_1}(\tfrac{t}{n}) = \{M_X(\tfrac{t}{n})\}^n = \left(E\left[e^{\frac{t}{n}X}\right]\right)^n \\ &\text{mgf of } s_2: && M_{s_2}(t) = \prod_{i=1}^{n} M_{X^2}(t) = \{M_{X^2}(t)\}^n = \left(E\left[e^{tX^2}\right]\right)^n \end{aligned} \tag{4.5}$$

We shall make use of these relations in the following examples.

⊕ ***Example 21:*** Sum of n Bernoulli Random Variables: The Binomial Distribution

Suppose that the discrete random variable X is Bernoulli distributed with parameter p. That is, $X \sim$ Bernoulli(p), where $P(X = 1) = p$, $P(X = 0) = 1 - p$, and $0 < p < 1$.

```
        g  =  p^x (1 - p)^(1-x);
domain[g]  =  {x, 0, 1} && {0 < p < 1} && {Discrete};
```

For a random sample of size n on X, the mgf of the sample sum s_1 is derived from (4.5) as:

```
mgf_s1 = Expect[e^(t x), g]^n
```

```
(1 + (-1 + e^t) p)^n
```

This is equivalent to the mgf of a Binomial(n, p) variable, as the reader can easily verify (use the *Discrete* palette to enter the Binomial pmf). Therefore, if $X \sim$ Bernoulli(p), then $s_1 \sim$ Binomial(n, p). ■

⊕ ***Example 22:*** Sum of n Exponential Random Variables: The Gamma Distribution

Let $X \sim$ Exponential(λ). For a random sample of size n, $(X_1, \ldots, X_n)$, we wish to find the distribution of the sample sum $s_1 = \sum_{i=1}^{n} X_i$.

Solution: Let $f(x)$ denote the pdf of X:

```
f = (1/λ) e^(-x/λ);        domain[f] = {x, 0, ∞} && {λ > 0};
```

By (4.5), the mgf of the sample sum s_1 is:

$$\mathbf{mgf_{s_1} = Expect[e^{t\,x}, f]^n}$$

$$\left(\frac{1}{1-t\,\lambda}\right)^n$$

This is identical to the mgf of a Gamma(a, b) random variable with parameter $a = n$, and $b = \lambda$, as we now verify:

$$\mathbf{g = \frac{x^{a-1}\, e^{-x/b}}{\Gamma[a]\, b^a};\quad domain[g] = \{x, 0, \infty\}\ \&\&\ \{a > 0, b > 0\};}$$

$$\mathbf{Expect[e^{t\,x}, g]}$$

$$(1 - b\,t)^{-a}$$

Thus, if $X \sim$ Exponential(λ), then $s_1 \sim$ Gamma(n, λ). ■

⊕ ***Example 23:*** Sum of n Chi-squared Random Variables

Let $X \sim \chi^2_v$, a Chi-squared random variable with v degrees of freedom, and let $(X_1, \ldots, X_n)$ denote a random sample of size n drawn from X. We wish to find the distribution of the sample sum $s_1 = \sum_{i=1}^n X_i$.

Solution: Let $f(x)$ denote the pdf of X:

$$\mathbf{f = \frac{x^{v/2-1}\, e^{-x/2}}{2^{v/2}\, \Gamma[\frac{v}{2}]};\qquad domain[f] = \{x, 0, \infty\}\ \&\&\ \{v > 0\};}$$

The mgf of X is:

$$\mathbf{mgf = Expect[e^{t\,x}, f]}$$

— This further assumes that: $\left\{t < \frac{1}{2}\right\}$

$$(1 - 2\,t)^{-v/2}$$

By (4.5), the mgf of the sample sum s_1 is:

$$\mathbf{mgf_{s_1} = mgf^n\ \ //\ PowerExpand}$$

$$(1 - 2\,t)^{-\frac{n\,v}{2}}$$

which is the mgf of a Chi-squared random variable with $n\,v$ degrees of freedom. Thus, if $X \sim \chi^2_v$, then $s_1 \sim \chi^2_{n\,v}$. ■

⊕ ***Example 24:*** Distribution of the Sample Mean for a Normal Random Variable

If $X \sim N(\mu, \sigma^2)$, find the distribution of the sample mean, for a random sample of size n.

Solution: Let $f(x)$ denote the pdf of X:

$$\mathbf{f} = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}; \quad \mathbf{domain[f] = \{x, -\infty, \infty\}\ \&\&\ \{\mu \in Reals, \sigma > 0\};}$$

Then the mgf of the sample mean, $\overline{X}$, is given by (4.5) as $\left(E\left[e^{\frac{t}{n}X}\right]\right)^n$:

$$\mathbf{Expect}\left[e^{\frac{t}{n}x}, \mathbf{f}\right]^n \quad \mathbf{//\ PowerExpand\ //\ Simplify}$$

$$e^{t\,\mu+\frac{t^2\,\sigma^2}{2\,n}}$$

which is the mgf of a $N\left(\mu, \frac{\sigma^2}{n}\right)$ variable. Therefore, $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. ■

⊕ ***Example 25:*** Distribution of the Sample Mean for a Cauchy Random Variable

Let X be a Cauchy random variable. We wish to find the distribution of the sample mean, $\overline{X}$, for a random sample of size n.

Solution: Let $f(x)$ denote the pdf of X:

$$\mathbf{f} = \frac{1}{\pi\,(1+x^2)}; \quad \mathbf{domain[f] = \{x, -\infty, \infty\};}$$

The mgf of a Cauchy random variable does not exist, so we shall use the characteristic function (cf) instead, as the latter always exists. Recall that the cf of X is $E[e^{i\,t\,X}]$:

$$\mathbf{cf = Expect}[e^{i\,t\,x}, \mathbf{f}]$$

```
— This further assumes that: {Im[t] == 0}
```

$$e^{-t\ \mathrm{Sign}[t]}$$

By (4.5), the cf of $\overline{X}$ is given by:

$$\mathbf{cf}_{\bar{X}} = \left(\mathbf{cf\ /.\ t} \to \frac{t}{n}\right)^n \quad \mathbf{//\ Simplify[\#, \{n > 0, n \in Integers\}]\ \&}$$

$$e^{-t\ \mathrm{Sign}[t]}$$

Note that the cf of $\overline{X}$ is identical to the cf of X. Therefore, if X is Cauchy, then $\overline{X}$ has the same distribution. ■

⊕ ***Example 26:*** Distribution of the Sample Sum of Squares for $X_i \sim N(\mu, 1)$
→ Derivation of a Noncentral Chi-squared Distribution

Let $(X_1, \ldots, X_n)$ be independent random variables, with $X_i \sim N(\mu, 1)$. We wish to find the density of the sample sum of squares $s_2 = \sum_{i=1}^n X_i^2$ using the mgf method.

Solution: Let $X \sim N(\mu, 1)$ have pdf $f(x)$:

$$\mathbf{f} = \frac{e^{-\frac{1}{2}(x-\mu)^2}}{\sqrt{2\pi}}; \qquad \mathbf{domain[f] = \{x, -\infty, \infty\};}$$

By (4.5), the mgf of s_2 is $(E[e^{t X^2}])^n$:

$$\mathbf{mgf = Expect}\left[e^{t x^2}, \mathbf{f}\right]^{n} \mathbf{// \; PowerExpand}$$

— This further assumes that: $\{t < \frac{1}{2}\}$

$$e^{\frac{n t \mu^2}{1-2t}} (1-2t)^{-n/2}$$

This expression is equivalent to the mgf of a noncentral Chi-squared variable $\chi_n^2(\lambda)$ with n degrees of freedom and noncentrality parameter $\lambda = n\mu^2$. To demonstrate this, we use **mathStatica**'s *Continuous* palette to input the $\chi_n^2(\lambda)$ pdf, and match its mgf to the one derived above:

$$\mathbf{f} = \frac{\mathbf{Hypergeometric0F1Regularized}[\frac{n}{2}, \frac{x\lambda}{4}]}{2^{n/2} e^{(x+\lambda)/2} x^{-(n-2)/2}};$$

$$\mathbf{domain[f] = \{x, 0, \infty\} \;\&\&\; \{n > 0, \lambda > 0\};}$$

Its mgf is given by:

$$\mathbf{Expect}[e^{t x}, \mathbf{f}]$$

— This further assumes that: $\{t < \frac{1}{2}\}$

$$e^{\frac{t\lambda}{1-2t}} (1-2t)^{-n/2}$$

We see that the mgf's are equivalent provided $\lambda = n\mu^2$, as claimed. Thus, if $X \sim N(\mu, 1)$, then $s_2 = \sum_{i=1}^n X_i^2 \sim \chi_n^2(n\mu^2)$. If $\mu = 0$, the noncentrality parameter disappears, and we revert to the familiar Chi-squared(n) pdf:

$$\mathbf{f \;/.\; \lambda \to 0}$$

$$\frac{2^{-n/2} e^{-x/2} x^{\frac{1}{2}(-2+n)}}{\Gamma[\frac{n}{2}]}$$

Figure 10 illustrates the noncentral Chi-squared pdf $\chi_{n=4}^2(\lambda)$, at different values of λ.

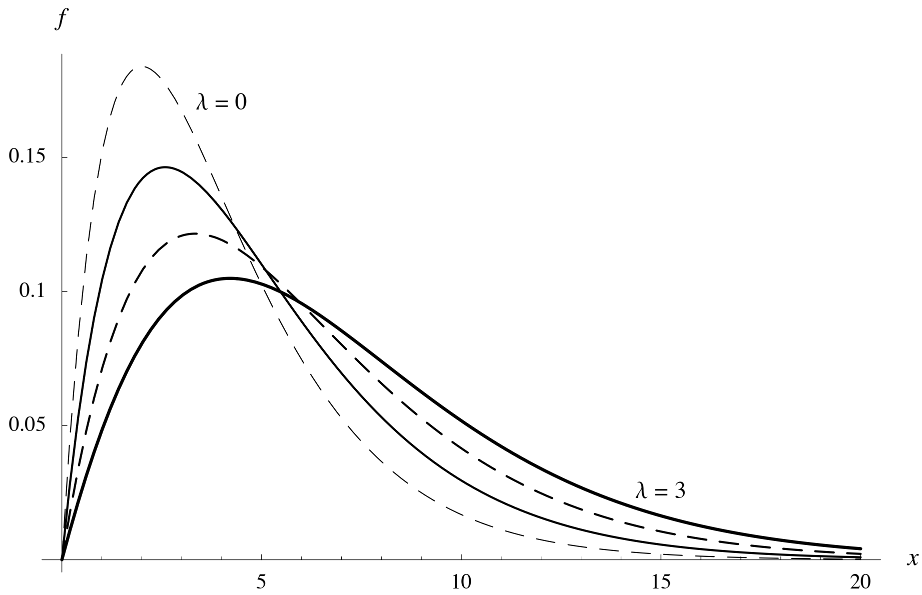


Fig. 10: Noncentral Chi-squared pdf when $n = 4$ and $\lambda = 0, 1, 2, 3$

⊕ ***Example 27:*** Distribution of the Sample Sum of Squares About the Mean

Let $(X_1, \ldots, X_n)$ be independent random variables, with $X_i \sim N(0, 1)$. We wish to find the density of the sum of squares about the sample mean; *i.e.* $SS = \sum_{i=1}^{n} (X_i - \overline{X})^2$ where $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$. Unlike previous examples, the random variable SS is not listed in (4.5). Nevertheless, we can find the solution by first applying a transformation known as *Helmert's transformation* and then applying a result obtained above with the mgf method. Helmert's transformation is given by:

$$\begin{aligned} Y_1 &= (X_1 - X_2) \big/ \sqrt{2} \\ Y_2 &= (X_1 + X_2 - 2X_3) \big/ \sqrt{6} \\ Y_3 &= (X_1 + X_2 + X_3 - 3X_4) \big/ \sqrt{12} \\ &\vdots \\ Y_{n-1} &= (X_1 + X_2 + \cdots + X_{n-1} - (n-1)X_n) \big/ \sqrt{n(n-1)} \\ Y_n &= (X_1 + X_2 + \cdots + X_n) \big/ \sqrt{n} \end{aligned} \tag{4.6}$$

For our purposes, the Helmert transformation has two important features:

(i) If each X_i is independent $N(0, 1)$, then each Y_i is also independent $N(0, 1)$.

(ii) $SS = \sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n-1} Y_i^2$.

The rest is easy: we know from *Example 26* that if $Y_i \sim N(0, 1)$, then $\sum_{i=1}^{n-1} Y_i^2$ is Chi-squared with $n - 1$ degrees of freedom. Therefore, for a random sample of size n on a standard Normal random variable, $\sum_{i=1}^{n} (X_i - \overline{X})^2 \sim \chi^2_{n-1}$.

To illustrate properties (i) and (ii), we can implement the Helmert transformation (4.6) in *Mathematica*:

$$\texttt{Helmert[n\_Integer] := Append}\Big[\texttt{Table}\Big[y_j == \frac{\sum_{i=1}^{j} x_i \;-\; j\, x_{j+1}}{\sqrt{j\,(j+1)}},\ \{j,\ n-1\}\Big],\ y_n == \frac{\sum_{i=1}^{n} x_i}{\sqrt{n}}\Big]$$

When, say, $n = 4$, we have:

$$\vec{X} = \texttt{Table}[x_i,\ \{i,\ 4\}];$$
$$\vec{Y} = \texttt{Table}[y_i,\ \{i,\ 4\}];$$
$$\texttt{eqn = Helmert[4]}$$

$$\Big\{\ y_1 == \frac{x_1 - x_2}{\sqrt{2}},\quad y_2 == \frac{x_1 + x_2 - 2\,x_3}{\sqrt{6}},\quad y_3 == \frac{x_1 + x_2 + x_3 - 3\,x_4}{2\,\sqrt{3}},\quad y_4 == \frac{1}{2}\,(x_1 + x_2 + x_3 + x_4)\ \Big\}$$

Let $f(\vec{x})$ denote the joint pdf of the X_i:

$$f = \prod_{i=1}^{n} \frac{e^{-\frac{x_i^2}{2}}}{\sqrt{2\,\pi}}\ \texttt{/. n} \to 4;\quad \texttt{domain[f] = Thread}\big[\{\vec{X},\ -\infty,\ \infty\}\big];$$

and let $g(\vec{y})$ denote the joint pdf of the Y_i:

$$g\ =\ \texttt{Transform[eqn, f]}$$
$$\texttt{domain[g]}\ =\ \texttt{Thread}\big[\{\vec{Y},\ -\infty,\ \infty\}\big];$$

$$\frac{e^{\frac{1}{2}\,(-y_1^2 - y_2^2 - y_3^2 - y_4^2)}}{4\,\pi^2}$$

Property (i) states that if the X_i are $N(0, 1)$, then the Y_i are also independent $N(0, 1)$. This is easily verified—the marginal distributions of each of Y_1, Y_2, Y_3 and Y_4:

$$\texttt{Map}\big[\ \texttt{Marginal[\#, g] \&},\ \vec{Y}\ \big]$$

$$\Big\{\frac{e^{-\frac{y_1^2}{2}}}{\sqrt{2\,\pi}},\ \frac{e^{-\frac{y_2^2}{2}}}{\sqrt{2\,\pi}},\ \frac{e^{-\frac{y_3^2}{2}}}{\sqrt{2\,\pi}},\ \frac{e^{-\frac{y_4^2}{2}}}{\sqrt{2\,\pi}}\Big\}$$

… are all $N(0, 1)$, while independence follows since the joint pdf $g(\vec{y})$ is equal to the product of the marginals.

Property (ii) states that $\sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n-1} Y_i^2$. To show this, we first find the inverse of the transformation equations:

```
inv = Solve[eqn, x⃗][[1]]
```

$$\left\{ x_4 \to \frac{1}{2}\left(-\sqrt{3}\, y_3 + y_4\right),\ x_3 \to \frac{1}{6}\left(-2\sqrt{6}\, y_2 + \sqrt{3}\, y_3 + 3\, y_4\right),\ x_1 \to \frac{1}{6}\left(3\sqrt{2}\, y_1 + \sqrt{6}\, y_2 + \sqrt{3}\, y_3 + 3\, y_4\right),\ x_2 \to \frac{1}{6}\left(-3\sqrt{2}\, y_1 + \sqrt{6}\, y_2 + \sqrt{3}\, y_3 + 3\, y_4\right) \right\}$$

and then examine the sum $\sum_{i=1}^{n} (X_i - \overline{X})^2$, given the transformation of X to Y:

$$\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \quad /.\ n \to 4 \quad /.\ \text{inv} \quad //\ \text{Simplify}$$

$$y_1^2 + y_2^2 + y_3^2$$

One final point is especially worth noting: since SS is a function of (Y_1, Y_2, Y_3), and since each of these variables is independent of Y_4, it follows that SS is independent of Y_4 or any function of it, including $Y_4 / \sqrt{n}$, which is equal to $\overline{X}$, by (4.6). Hence, in Normal samples, SS is independent of $\overline{X}$. This applies not only when $n = 4$, but also quite generally for arbitrary n. The independence of SS and $\overline{X}$ in Normal samples is an important property that is useful when constructing statistics for hypothesis testing. ■

4.6 Exercises

1. Let $X \sim$ Uniform(0, 1). Show that the distribution of $Y = \log\left(\frac{X}{1-X}\right)$ is standard Logistic.

2. Let $X \sim N(\mu, \sigma^2)$. Find the distribution of $Y = \exp(\exp(X))$.

3. Find the pdf of $Y = 1/X$:

(i) if $X \sim$ Gamma(a, b); (Y has an InverseGamma(a, b) distribution).

(ii) if $X \sim$ PowerFunction(a, c); (Y has a Pareto distribution).

(iii) if $X \sim$ InverseGaussian(μ, λ); (Y has a Random Walk distribution).
Plot the Random Walk pdf when $\mu = 1$ and $\lambda = 1$, 4 and 16.

4. Let X have a Maxwell–Boltzmann distribution. Find the distribution of $Y = X^2$ using both the Transformation Method and the MGF Method.

5. Let X_1 and X_2 have joint pdf $f(x_1, x_2) = 4\, x_1 x_2$, $0 < x_1 < 1$, $0 < x_2 < 1$. Find the joint pdf of $Y_1 = X_1^2$ and $Y_2 = X_1 X_2$. Plot the domain of support of Y_1 and Y_2.

6. Let X_1 and X_2 be independent standard Cauchy random variables. Find the distribution of $Y = X_1 X_2$ and plot it.

7. Let X_1 and X_2 be independent Gamma variates with the same scale parameter b. Find the distribution of $Y = \frac{X_1}{X_1 + X_2}$.

8. Let $X \sim \text{Geometric}(p)$ and $Y \sim \text{Geometric}(q)$ be independent random variables. Find the distribution of $Z = Y - X$. Plot the pmf of Z when (i) $p = q = \frac{1}{2}$, (ii) $p = \frac{1}{2}$, $q = \frac{1}{8}$.

9. Find the sum of n independent Gamma(a, b) random variables.