

Chapter 2

Quadratic Forms of Random Variables

2.1 Quadratic Forms

For a $k \times k$ symmetric matrix $\mathbf{A} = \{a_{ij}\}$ the quadratic function of k variables $\mathbf{x} = (x_1, \dots, x_n)'$ defined by

$$Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^k \sum_{j=1}^k a_{i,j}x_i x_j$$

is called the *quadratic form* with matrix \mathbf{A} .

If \mathbf{A} is not symmetric, we can have an equivalent expression/quadratic form replacing \mathbf{A} by $(\mathbf{A} + \mathbf{A}')/2$.

Definition 1. $Q(\mathbf{x})$ and the matrix \mathbf{A} are called positive definite if

$$Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} > 0, \quad \forall \mathbf{x} \in \mathbb{R}^k, \mathbf{x} \neq \mathbf{0}$$

and positive semi-definite if

$$Q(\mathbf{x}) \geq \forall \mathbf{x} \in \mathbb{R}^k$$

For negative definite and negative semi-definite, replace the $>$ and \geq in the above definitions by $<$ and \leq , respectively.

Theorem 1. A symmetric matrix \mathbf{A} is positive definite if and only if it has a Cholesky decomposition $\mathbf{A} = \mathbf{R}'\mathbf{R}$ with strictly positive diagonal elements in \mathbf{R} , so that \mathbf{R}^{-1} exists. (In practice this means that none of the diagonal elements of \mathbf{R} are very close to zero.)

Proof. The “if” part is proven by construction. The Cholesky decomposition, \mathbf{R} , is constructed a row at a time and the diagonal elements are evaluated as the square roots of expressions calculated from the current row of \mathbf{A} and previous rows of \mathbf{R} . If the expression whose square root is to be calculated is not positive then you can determine a non-zero $\mathbf{x} \in \mathbb{R}^k$ for which $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 0$.

Suppose that $\mathbf{A} = \mathbf{R}'\mathbf{R}$ with \mathbf{R} invertible. Then

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{R}'\mathbf{R}\mathbf{x} = \|\mathbf{R}\mathbf{x}\|^2 \geq 0$$

with equality only if $\mathbf{R}\mathbf{x} = \mathbf{0}$. But if \mathbf{R}^{-1} exists then $\mathbf{x} = \mathbf{R}^{-1}\mathbf{0}$ must also be zero. □

Transformation of Quadratic Forms:

Theorem 2. Suppose that \mathbf{B} is a $k \times k$ nonsingular matrix. Then the quadratic form $Q^*(\mathbf{y}) = \mathbf{y}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{y}$ is positive definite if and only if $Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ is positive definite. Similar results hold for positive semi-definite, negative definite and negative semi-definite.

Proof.

$$Q^*(\mathbf{y}) = \mathbf{y}'\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{y} = \mathbf{x}'\mathbf{A}\mathbf{x} > 0$$

where $\mathbf{x} = \mathbf{B}\mathbf{y} \neq \mathbf{0}$ because $\mathbf{y} \neq \mathbf{0}$ and \mathbf{B} is nonsingular. \square

Theorem 3. For any $k \times k$ symmetric matrix \mathbf{A} the quadratic form defined by \mathbf{A} can be written using its spectral decomposition as

$$Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^k \lambda_i \|\mathbf{q}'_i \mathbf{x}\|^2$$

where the eigendecomposition of \mathbf{A} is $\mathbf{Q}'\mathbf{\Lambda}\mathbf{Q}$ with $\mathbf{\Lambda}$ diagonal with diagonal elements λ_i , $i = 1, \dots, k$, \mathbf{Q} is the orthogonal matrix with the eigenvectors, \mathbf{q}_i , $i = 1, \dots, k$ as its columns. (Be careful to distinguish the bold face \mathbf{Q} , which is a matrix, from the unbolded $Q(\mathbf{x})$, which is the quadratic form.)

Proof. For any $\mathbf{x} \in \mathbb{R}^k$ let $\mathbf{y} = \mathbf{Q}'\mathbf{x} = \mathbf{Q}^{-1}\mathbf{x}$. Then

$$Q(\mathbf{x}) = \text{tr}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \text{tr}(\mathbf{x}'\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'\mathbf{x}) = \text{tr}(\mathbf{y}'\mathbf{\Lambda}\mathbf{y}) = \text{tr}(\mathbf{\Lambda}\mathbf{y}\mathbf{y}') = \sum_{i=1}^k \lambda_i y_i^2 = \sum_{i=1}^k \lambda_i \|\mathbf{q}'_i \mathbf{x}\|^2$$

This proof uses a common “trick” of expressing the scalar $Q(\mathbf{x})$ as the trace of a 1×1 matrix so we can reverse the order of some matrix multiplications. \square

Corollary 1. A symmetric matrix \mathbf{A} is positive definite if and only if its eigenvalues are all positive, negative definite if and only if its eigenvalues are all negative, and positive semi-definite if all its eigenvalues are non-negative.

Corollary 2. $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{\Lambda})$ hence $\text{rank}(\mathbf{A})$ equals the number of non-zero eigenvalues of \mathbf{A}

2.2 Idempotent Matrices

Definition 2 (Idempotent). The $k \times k$ matrix \mathbf{A} , is idempotent if $\mathbf{A}^2 = \mathbf{A}\mathbf{A} = \mathbf{A}$.

Definition 3 (Projection matrices). A symmetric, idempotent matrix \mathbf{A} is a projection matrix. The effect of the mapping $\mathbf{x} \rightarrow \mathbf{A}\mathbf{x}$ is orthogonal projection of \mathbf{x} onto $\text{col}(\mathbf{A})$.

Theorem 4. All the eigenvalues of an idempotent matrix are either zero or one.

Proof. Suppose that λ is an eigenvalue of the idempotent matrix \mathbf{A} . Then there exists a non-zero \mathbf{x} such that $\mathbf{Ax} = \lambda\mathbf{x}$. But $\mathbf{Ax} = \mathbf{AAx}$ because \mathbf{A} is idempotent. Thus

$$\lambda\mathbf{x} = \mathbf{Ax} = \mathbf{AAx} = \mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{Ax}) = \lambda^2\mathbf{x}$$

and

$$\mathbf{0} = \lambda^2\mathbf{x} - \lambda\mathbf{x} = \lambda(\lambda - 1)\mathbf{x}$$

for some non-zero \mathbf{x} , which implies that $\lambda = 0$ or $\lambda = 1$. \square

Corollary 3. *The $k \times k$ symmetric matrix \mathbf{A} is idempotent of rank(\mathbf{A}) = r iff \mathbf{A} has r eigenvalues equal to 1 and $k - r$ eigenvalues equal to 0*

Proof. A matrix \mathbf{A} with r eigenvalues of 1 and $k - r$ eigenvalues of zero has r non-zero eigenvalues and hence rank(\mathbf{A}) = r . Because \mathbf{A} is symmetric its eigendecomposition is $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$ for an orthogonal \mathbf{Q} and a diagonal $\mathbf{\Lambda}$. Because the eigenvalues of $\mathbf{\Lambda}$ are the same as those of \mathbf{A} , they must be all zeros or ones. That is all the diagonal elements of $\mathbf{\Lambda}$ are zero or one. Hence $\mathbf{\Lambda}$ is idempotent, $\mathbf{\Lambda}\mathbf{\Lambda} = \mathbf{\Lambda}$, and

$$\mathbf{AA} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}' = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}' = \mathbf{A}$$

is also idempotent. \square

Corollary 4. *For a symmetric idempotent matrix \mathbf{A} , we have $\text{tr}(\mathbf{A}) = \text{rank}(\mathbf{A})$, which is the dimension of $\text{col}(\mathbf{A})$, the space into which \mathbf{A} projects.*

2.3 Expected Values and Covariance Matrices of Random Vectors

An k -dimensional *vector-valued random variable* (or, more simply, a *random vector*), \mathcal{X} , is a k -vector composed of k scalar random variables

$$\mathcal{X} = (\mathcal{X}_1, \dots, \mathcal{X}_k)'$$

If the expected values of the component random variables are $\mu_i = E(\mathcal{X}_i)$, $i = 1, \dots, k$ then

$$E(\mathcal{X}) = \boldsymbol{\mu}_{\mathcal{X}} = (\mu_1, \dots, \mu_k)'$$

Suppose that $\mathcal{Y} = (\mathcal{Y}_1, \dots, \mathcal{Y}_m)'$ is an m -dimensional random vector, then the *covariance* of \mathcal{X} and \mathcal{Y} , written $\text{Cov}(\mathcal{X}, \mathcal{Y})$ is

$$\boldsymbol{\Sigma}_{\mathcal{XY}} = \text{Cov}(\mathcal{X}, \mathcal{Y}) = E[(\mathcal{X} - \boldsymbol{\mu}_{\mathcal{X}})(\mathcal{Y} - \boldsymbol{\mu}_{\mathcal{Y}})']$$

The *variance-covariance* matrix of \mathcal{X} is

$$\text{Var}(\mathcal{X}) = \boldsymbol{\Sigma}_{\mathcal{XX}} = E[(\mathcal{X} - \boldsymbol{\mu}_{\mathcal{X}})(\mathcal{X} - \boldsymbol{\mu}_{\mathcal{X}})']$$

Suppose that \mathbf{c} is a constant m -vector, \mathbf{A} is a constant $m \times k$ matrix and $\mathcal{Z} = \mathbf{AZ} + \mathbf{c}$ is a linear transformation of \mathcal{X} . Then

$$E(\mathcal{Z}) = \mathbf{A}E(\mathcal{X}) + \mathbf{c}$$

and

$$\text{Var}(\mathcal{Z}) = \mathbf{A} \text{Var}(\mathcal{X}) \mathbf{A}'$$

If we let $\mathcal{W} = \mathbf{B}\mathcal{Y} + \mathbf{d}$ be a linear transformation of \mathcal{Y} for suitably sized \mathbf{B} and \mathbf{d} then

$$\text{Cov}(\mathcal{Z}, \mathcal{W}) = \mathbf{A} \text{Cov}(\mathcal{X}, \mathcal{Y}) \mathbf{B}'$$

Theorem 5. *The variance-covariance matrix $\Sigma_{\mathcal{X}, \mathcal{X}}$ of \mathcal{X} is a symmetric and positive semi-definite matrix*

Proof. The result follows from the property that the variance of a scalar random variable is non-negative. Suppose that \mathbf{b} is any nonzero, constant k -vector. Then

$$0 \leq \text{Var}(\mathbf{b}'\mathcal{X}) = \mathbf{b}'\Sigma_{\mathcal{X}, \mathcal{X}}\mathbf{b}$$

which is the positive, semi-definite condition. □

2.4 Mean and Variance of Quadratic Forms

Theorem 6. *Let \mathcal{X} be a k -dimensional random vector and \mathbf{A} be a constant $k \times k$ symmetric matrix. If $E(\mathcal{X}) = \boldsymbol{\mu}$ and $\text{Var}(\mathcal{X}) = \Sigma$, then*

$$E(\mathcal{X}'\mathbf{A}\mathcal{X}) = \text{tr}(\mathbf{A}\Sigma) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$$

Proof.

$$\begin{aligned} E(\mathcal{X}'\mathbf{A}\mathcal{X}) &= \text{tr}(E(\mathcal{X}'\mathbf{A}\mathcal{X})) \\ &= E[\text{tr}(\mathcal{X}'\mathbf{A}\mathcal{X})] \\ &= E[\text{tr}(\mathbf{A}\mathcal{X}\mathcal{X}')] \\ &= \text{tr}(\mathbf{A}E[\mathcal{X}\mathcal{X}']) \\ &= \text{tr}(\mathbf{A}(\text{Cov}(\mathcal{X}) + \boldsymbol{\mu}\boldsymbol{\mu}')) \\ &= \text{tr}(\mathbf{A}\Sigma_{\mathcal{X}, \mathcal{X}}) + \text{tr}(\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}') \\ &= \text{tr}(\mathbf{A}\Sigma_{\mathcal{X}, \mathcal{X}}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} \end{aligned}$$

□

2.5 Distribution of Quadratic Forms in Normal Random Variables

Definition 4 (Non-Central χ^2). *If \mathcal{X} is a (scalar) normal random variable with $E(\mathcal{X}) = \mu$ and $\text{Var}(\mathcal{X}) = 1$, then the random variable $\mathcal{V} = \mathcal{X}^2$ is distributed as $\chi_1^2(\lambda^2)$, which is called the noncentral χ^2 distribution with 1 degree of freedom and non-centrality parameter $\lambda^2 = \mu^2$. The mean and variance of \mathcal{V} are*

$$E[\mathcal{V}] = 1 + \lambda^2 \text{ and } \text{Var}[\mathcal{V}] = 2 + 4\lambda^2$$

As described in the previous chapter, we are particularly interested in random n -vectors, \mathbf{Y} , that have a *spherical normal distribution*.

Theorem 7. Let $\mathcal{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$ be an n -vector with a spherical normal distribution and \mathbf{A} be an $n \times n$ symmetric matrix. Then the ratio $\mathcal{Y}'\mathbf{A}\mathcal{Y}/\sigma^2$ will have a $\chi_r^2(\lambda^2)$ distribution with $\lambda^2 = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/\sigma^2$ if and only if \mathbf{A} is idempotent with $\text{rank}(\mathbf{A}) = r$

Proof. Suppose that \mathbf{A} is idempotent (which, in combination with being symmetric, means that it is a projection matrix) and has $\text{rank}(\mathbf{A}) = r$. Its eigendecomposition, $\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}'$, is such that \mathbf{V} is orthogonal and $\boldsymbol{\Lambda}$ is $n \times n$ diagonal with exactly $r = \text{rank}(\mathbf{A})$ ones and $n - r$ zeros on the diagonal. Without loss of generality we can (and do) arrange the eigenvalues in decreasing order so that $\lambda_j = 1, j = 1, \dots, r$ and $\lambda_j = 0, j = r + 1, \dots, n$. Let $\mathcal{X} = \mathbf{V}'\mathcal{Y}$

$$\begin{aligned} \frac{\mathcal{Y}'\mathbf{A}\mathcal{Y}}{\sigma^2} &= \frac{\mathcal{Y}'\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}'\mathcal{Y}}{\sigma^2} \\ &= \frac{\mathcal{X}'\boldsymbol{\Lambda}\mathcal{X}}{\sigma^2} \\ &= \sum_{j=1}^n \lambda_j \frac{\mathcal{X}_j^2}{\sigma^2} \\ &= \sum_{j=1}^r \frac{\mathcal{X}_j^2}{\sigma^2} \end{aligned}$$

(Notice that the last sum is to $j = r$, not $j = n$.) However, $\frac{\mathcal{X}_j}{\sigma} \sim \mathcal{N}(\mathbf{v}_j'\boldsymbol{\mu}/\sigma, 1)$ so $\frac{\mathcal{X}_j^2}{\sigma^2} \sim \chi_1^2((\mathbf{v}_j'\boldsymbol{\mu}/\sigma)^2)$. Therefore

$$\sum_{j=1}^r \frac{\mathcal{X}_j^2}{\sigma^2} \sim \chi_{(r)}^2(\lambda^2) \text{ where } \lambda^2 = \frac{\boldsymbol{\mu}'\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}'\boldsymbol{\mu}}{\sigma^2} = \frac{\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}}{\sigma^2}$$

□

Corollary 5. For \mathbf{A} a projection of rank r , $(\mathcal{Y}'\mathbf{A}\mathcal{Y})/\sigma^2$ has a central χ_r^2 distribution if and only if $\mathbf{A}\boldsymbol{\mu} = \mathbf{0}$

Proof. The χ_r^2 distribution will be central if and only if

$$0 = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{A}\mathbf{A}\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{A}'\mathbf{A}\boldsymbol{\mu} = \|\mathbf{A}\boldsymbol{\mu}\|^2$$

□

Corollary 6. In the full-rank Gaussian linear model, $\mathbf{Y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n)$, the residual sum of squares, $\|\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2$ has a central $\sigma^2 \chi_{n-r}^2$ distribution.

Proof. In the full rank model with the QR decomposition of \mathbf{X} given by

$$\mathbf{X} = [\mathbf{Q}_1 \quad \mathbf{Q}_2] \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$$

and \mathbf{R} invertible, the fitted values are $\mathbf{Q}_1\mathbf{Q}'_1\mathcal{Y}$ and the residuals are $\mathbf{Q}_2\mathbf{Q}'_2\mathbf{y}$ so the residual sum of squares is the quadratic form $\mathcal{Y}'\mathbf{Q}_2\mathbf{Q}'_2\mathcal{Y}$. The matrix defining the quadratic form, $\mathbf{Q}_2\mathbf{Q}'_2$, is a projection matrix. It is obviously symmetric and it is idempotent because $\mathbf{Q}_2\mathbf{Q}'_2\mathbf{Q}_2\mathbf{Q}'_2 = \mathbf{Q}_2\mathbf{Q}'_2$. As

$$\mathbf{Q}'_2\boldsymbol{\mu} = \mathbf{Q}'_2\mathbf{X}\boldsymbol{\beta}_0 = \mathbf{Q}'_2\mathbf{Q}_1\mathbf{R}\boldsymbol{\beta}_0 = \underbrace{\mathbf{0}}_{(n-p)\times n} \mathbf{R}\boldsymbol{\beta}_0 = \underbrace{\mathbf{0}}_{(n-p)\times p} \boldsymbol{\beta}_0 = \mathbf{0}_{n-p}$$

the ratio

$$\frac{\mathcal{Y}'\mathbf{Q}_2\mathbf{Q}'_2\mathcal{Y}}{\sigma^2} \sim \chi^2_{n-p}$$

and the RSS has a central $\sigma^2\chi^2_{n-p}$ distribution. □

R Exercises: Let's check some of these results by simulation. First we claim that if $\mathcal{X} \sim \mathcal{N}(\mu, 1)$ then $\mathcal{X}^2 \sim \chi^2(\lambda^2)$ where $\lambda^2 = \mu^2$. First simulate from a standard normal distribution

```
> set.seed(1234) # reproducible "random" values
> X <- rnorm(100000) # standard normal values
> zapsmall(summary(V <- X^2)) # a very skew distribution
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000  0.1026  0.4521  0.9989  1.3190 20.3300
```

```
> var(V)
```

```
[1] 1.992403
```

The mean and variance of the simulated values agree quite well with the theoretical values of 1 and 2, respectively.

To check the form of the distribution we could plot an empirical density function but this distribution has its maximum density at 0 and is zero to the left of 0 so an empirical density is a poor indication of the actual shape of the density. Instead, in Fig. 2.1, we present the quantile-quantile plot for this sample versus the (theoretical) quantiles of the χ^2_1 distribution.

Now simulate a non-central χ^2 with non-centrality parameter $\lambda^2 = 4$

```
> V1 <- rnorm(100000, mean=2)^2
> zapsmall(summary(V1))
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.000  1.773  3.994  5.003  7.144 39.050
```

```
> var(V1)
```

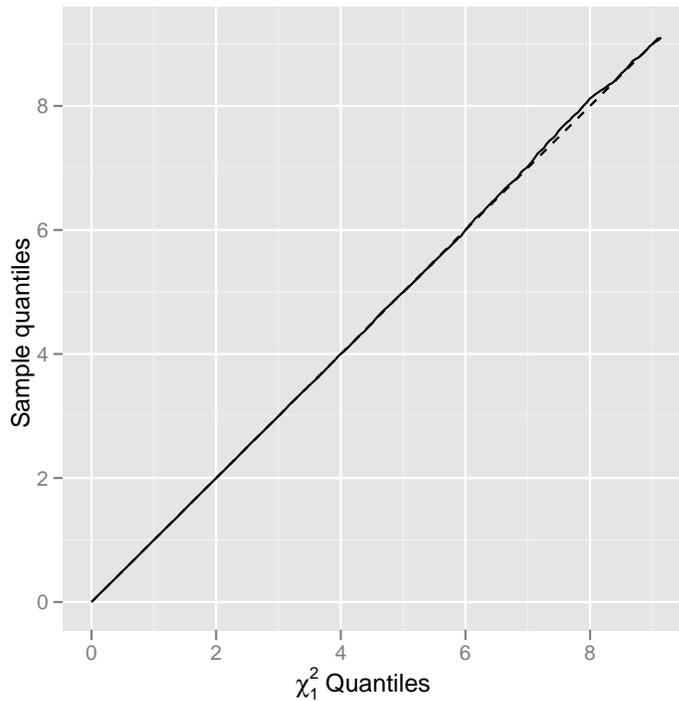


Figure 2.1: A quantile-quantile plot of the squares of simulated $\mathcal{N}(0, 1)$ random variables versus the quantiles of the χ_1^2 distribution. The dashed line is a reference line through the origin with a slope of 1.

```
[1] 17.95924
```

The sample mean is close to the theoretical value of $5 = 1 + \lambda^2$ and the sample variance is close to the theoretical value of $2 + 4\lambda^2$ although perhaps not as close as one would hope in a sample of size 100,000.

A quantile-quantile plot versus the non-central distribution, $\chi_1^2(4)$, (Fig. 2.2) and versus the central distribution, χ_1^2 , shows that the sample does follow the claimed distribution $\chi_1^2(4)$ and is stochastically larger than the χ_1^2 distribution.

More interesting, perhaps is the distribution of the residual sum of squares from a regression model. We simulate from our previously fitted model `lm1`

```
> lm1 <- lm(optden ~ carb, Formaldehyde)
> str(Ymat <- data.matrix(unname(simulate(lm1, 10000))))

num [1:6, 1:10000] 0.088 0.258 0.444 0.521 0.619 ...
- attr(*, "dimnames")=List of 2
..$ : chr [1:6] "1" "2" "3" "4" ...
..$ : NULL
```

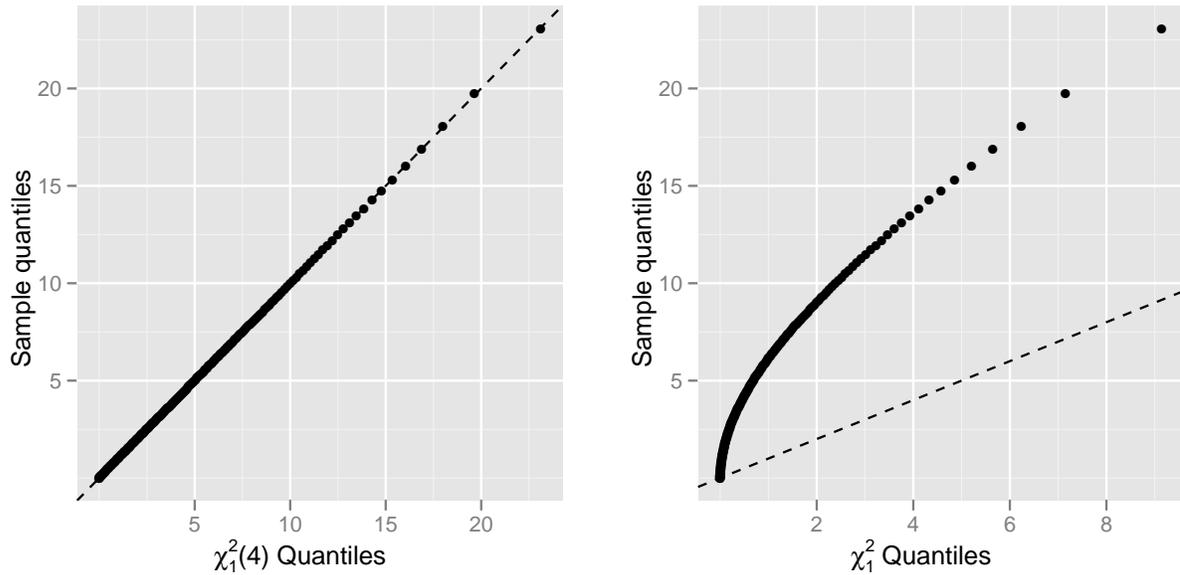


Figure 2.2: Quantile-quantile plots of a sample of squares of $\mathcal{N}(2, 1)$ random variables versus the quantiles of a $\chi_1^2(4)$ non-central distribution (left panel) and a χ_1^2 central distribution (right panel)

```
> str(RSS <- deviance(fits <- lm(Ymat ~ carb, Formaldehyde)))
num [1:10000] 0.000104 0.000547 0.00055 0.000429 0.000228 ...
> fits[["df.residual"]]
[1] 4
```

Here the `Ymat` matrix is 10,000 simulated response vectors from model `lm1` using the estimated parameters as the true values of β and σ^2 . Notice that we can fit the model to **all** 10,000 response vectors in a single call to the `lm()` function.

The `deviance()` function applied to a model fit by `lm()` returns the residual sum of square, which is not technically the deviance but is often the quantity of interest.

These simulated residual sums of squares should have a $\sigma^2 \chi_4^2$ distribution where σ^2 is the residual sum of squares in model `lm1` divided by 4.

```
> (sigsq <- deviance(lm1)/4)
[1] 7.48e-05
> summary(RSS)
      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
6.997e-07 1.461e-04 2.537e-04 3.026e-04 4.114e-04 1.675e-03
```

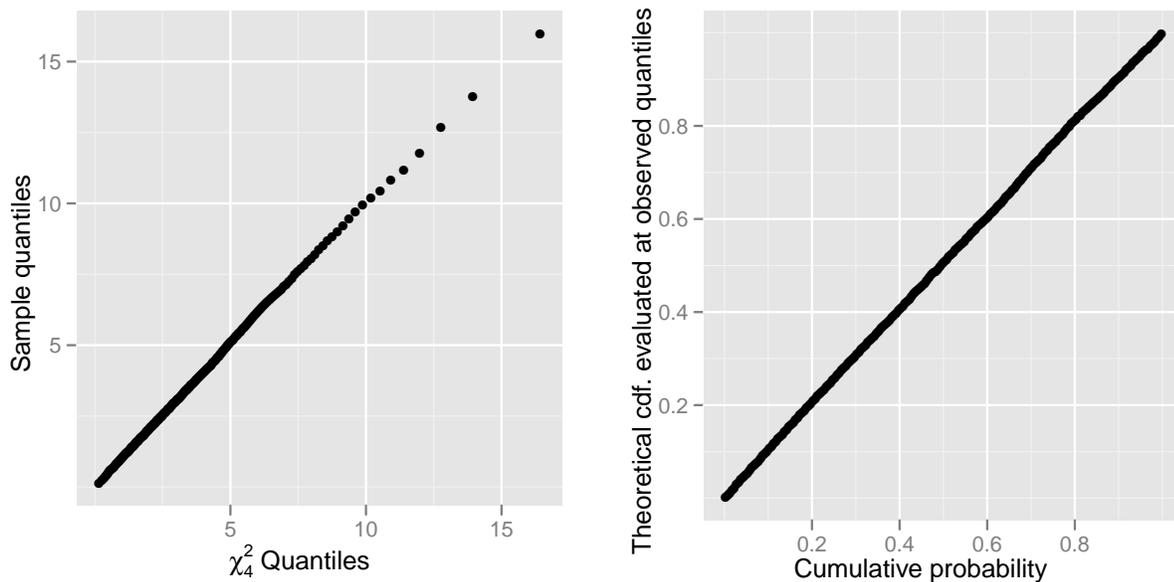


Figure 2.3: Quantile-quantile plot of the scaled residual sum of squares, RSS_{sq} , from simulated responses versus the quantiles of a χ_4^2 distribution (left panel) and the corresponding probability-probability plot on the right panel.

We expect a mean of $4\sigma^2$ and a variance of $2 \cdot 4(\sigma^2)^2$. It is easier to see this if we divide these values by σ^2

```
> summary(RSSsc <- RSS/sigsq)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.009354	1.953000	3.392000	4.045000	5.500000	22.390000

```
> var(RSSsc)
```

```
[1] 8.000057
```

A quantile-quantile plot with respect to the χ_4^2 distribution (Fig. 2.3) shows very good agreement between the empirical and theoretical quantiles. Also shown in Fig. 2.3 is the probability-probability plot. Instead of plotting the sample quantiles versus the theoretical quantiles we take equally spaced values on the probability scale (function `ppoints()`), evaluate the sample quantiles and then apply the theoretical cdf to the empirical quantiles. This should also produce a straight line. It has the advantage that the points are equally spaced on the x-axis.

We could also plot the empirical density of these simulated values and overlay it with the theoretical density (Fig. 2.4).

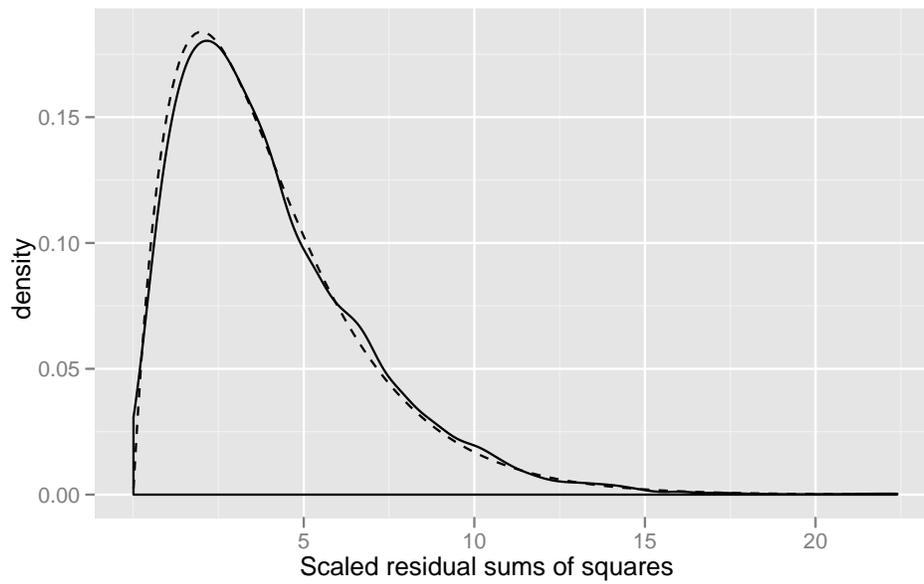


Figure 2.4: Empirical density plot of the scaled residual sums of squares, `RSSsq`, from simulated responses. The overlaid dashed line is the density of a χ_4^2 random variable. The peak of the empirical density gets shifted a bit to the right because of the way the empirical density is calculated. It uses a symmetric kernel which is not a good choice for a skewed density like this.