Lecture 4: Bivariate Distributions

Today's lecture

In this lecture we consider two related topics:

1. General simulation from a *bivariate distribution*
2. Simulation from the *bivariate Normal distribution*

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Assignment:
  - Deadline: This coming Thursday, 3pm
  - Hard-copy submission, with attached NESS cover sheet, to the General Office
  - R code: only need to include where it's obvious that you'll get marks for it
  - Graphs: include when asked, and always provide comments
  - Question 1: see practical sheet 1; Question 2: see chapter 2 and practical sheet 2; Question 3: see chapter 3 and practical sheet 3
  - Time for last-minute help in this week's practical session (Thurs @ 11)

Some other stuff first...

- Test:
  - Next Tuesday, 9am Herschel PC cluster
  - Strike action: Test should still go ahead
  - Mock test now available in Blackboard
  - Revision session this coming Friday, 1pm LT2

Some other stuff first...

- Test:
  - Next Tuesday, 9am Herschel PC cluster
  - Strike action: Test should still go ahead
  - Mock test now available in Blackboard
  - Revision session this coming Friday, 1pm LT2

Some other stuff first...

- Test:
  - Next Tuesday, 9am Herschel PC cluster
  - Strike action: Test should still go ahead
  - Mock test now available in Blackboard
  - Revision session this coming Friday, 1pm LT2

Some other stuff first...

- Test:
  - Next Tuesday, 9am Herschel PC cluster
  - Strike action: Test should still go ahead
  - Mock test now available in Blackboard
  - Revision session this coming Friday, 1pm LT2

Some other stuff first...

- Test:
  - Next Tuesday, 9am Herschel PC cluster
  - Strike action: Test should still go ahead
  - Mock test now available in Blackboard
  - Revision session this coming Friday, 1pm LT2

Some other stuff first...

- Test:
  - Next Tuesday, 9am Herschel PC cluster
  - Strike action: Test should still go ahead
  - Mock test now available in Blackboard
  - Revision session this coming Friday, 1pm LT2

Simulation using a marginal and conditional

Suppose we have a pair of continuous random variables (X, Y) with joint PDF $f_{X,Y}(x, y)$.

Recall that

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X}(y \mid X = x)$$

where $f_X(x)$ is the marginal PDF of X and $f_{Y|X}(y \mid X = x)$ is the conditional PDF of Y given $X = x$.

One way to simulate realizations of the pair (X, Y) is:

1. Simulate a realization x of X from the marginal distribution of X.
2. Simulate a realization y of Y from the conditional distribution of Y given $X = x$.

Simulation using a marginal and conditional

Example

Suppose $X \sim U(0.1, 0.5)$ and

$$f_{Y|X}(Y \mid X = x) = \begin{cases} xe^{-xy}, & \text{when } y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Write an R function to generate a $2 \times n$ matrix of samples from this distribution, and produce a scatter plot.

Example 4.1 – R code

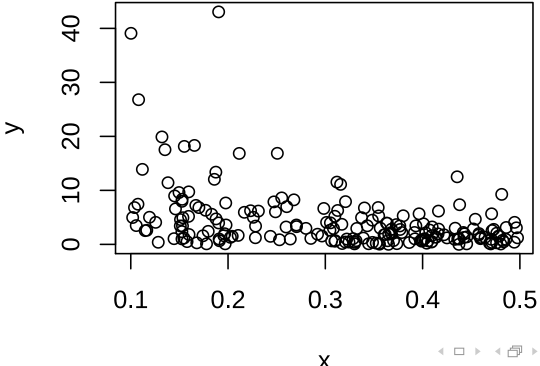
```
example4.1 = function(n) {
    output = matrix(0, 2, n)
    x = runif(n, 0.1, 0.5)
    y = rexp(n, rate = x)
    output[1, ] = x
    output[2, ] = y
    output
}
```

Make some notes about the R function.

Example 4.1 – plot

To obtain a scatterplot:

```
test = example4.1(200)
plot(output[1,], output[2,], xlab = 'X', ylab = 'Y')
```



The trivariate case

Example

Suppose that $X \sim U(-1,1)$ and $Y \sim Exp(\lambda)$ independently of X and that

$$f_{Z|X,Y}(Z \mid X = x, Y = y) = \frac{1}{\sqrt{2\pi}y} e^{-\frac{1}{2}\left(\frac{z-x}{y}\right)^2}.$$

Write an R function to sample realizations (X, Y, Z) in the case $\lambda = 5$ and produce a scatterplot of (X, Z).
Note that, given that $X = x$, $Y = y$ then Z has a normal distribution:

$$Z \mid X = x, Y = y \sim N(x, y^2).$$

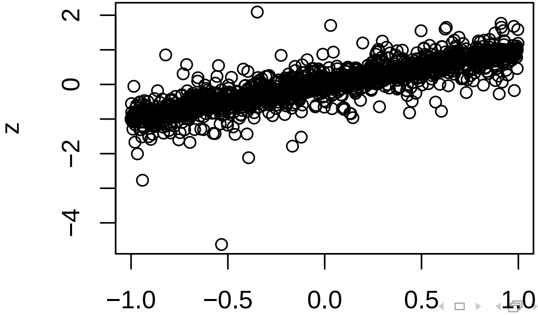
Example 4.2 – R function

```
example4.2 = function(n) {
   output = matrix(0, 3, n)
    x = runif(n, -1, 1)
    y = rexp(n, rate = 5)
    z = rnorm(n, mean = x, sd = y)
    output[1, ] = x
    output[2, ] = y
    output[3, ] = z
    output
}
```

Make some notes about the R function.

Example 4.2 – scatterplot

```
test = example4.2(2000)
plot(test[1, ], test[3, ], xlab = 'x', ylab = 'z')
```



The multivariate Normal distribution

You met the bivariate Normal distribution in MAS1604.

The bivariate Normal distribution is a generalization of the Normal distribution to pairs of random variables, or equivalently, to a distribution on vectors in $\mathbb{R}^2$.

The multivariate Normal distribution is the analagous distribution to vectors in $\mathbb{R}^n$.

The bivariate Normal distribution

Suppose μ_x, μ_y, $\sigma_x \geq 0$, $\sigma_y \geq 0$ and $-1 \leq \rho \leq 1$ are constants. Define the 2×2 matrix Σ by

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}.$$

Then define a joint probability density function by

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}Q(x,y)\right)$$

where

$$Q(x,y) = (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})$$

and

$$\underline{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \underline{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}.$$

If random variables (X, Y) have joint probability density given by $f_{X,Y}$ above, then we say that (X, Y) have a bivariate normal distribution and write

$$(X, Y)^T \sim N_2(\underline{\mu}, \Sigma).$$

It can be proved that the function $f_{X,Y}(x, y)$ integrates to 1 and therefore defines a valid joint pdf.

The notes contain expansions of $Q(x, y)$ and $\det \Sigma$.

Remarks

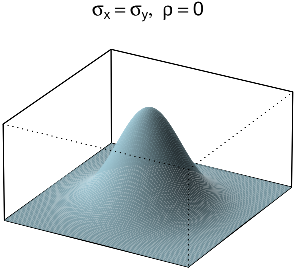
1. The vector $\underline{\mu} = (\mu_x, \mu_y)^T$ is called the mean vector and the matrix Σ is called the covariance matrix (or sometimes variance-covariance matrix).
2. Functions of the form $F(\underline{x}) = \underline{x}^T \Sigma^{-1} \underline{x}$ are called quadratic forms. Quadratic forms are functions $\mathbb{R}^n \to \mathbb{R}$ which satisfy certain properties. They crop up in several areas of mathematics and statistics.
3. The matrix Σ and its inverse Σ^{-1} are positive definite. A matrix A is positive definite if

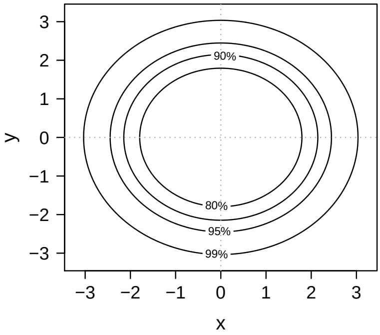
$$\underline{x}^T A \underline{x} \geq 0$$

for all non-zero vectors $\underline{x}$.

4. It follows that when $\mu_x = \mu_y = 0$, $Q(x, y)$ is a positive definite quadratic form.

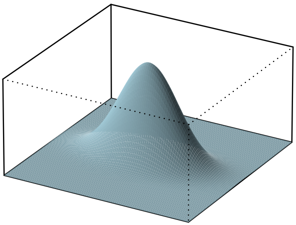
Pictures

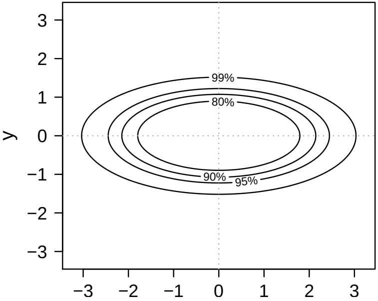




Pictures

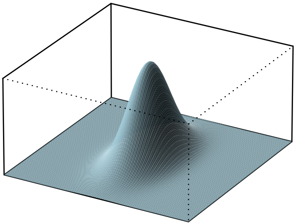
$\sigma_x = 2\sigma_y,\ \rho = 0$

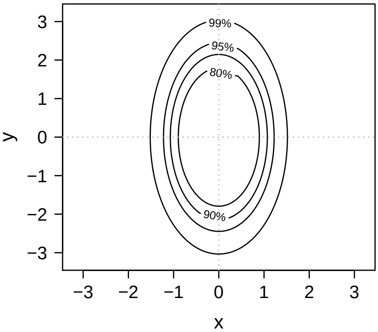




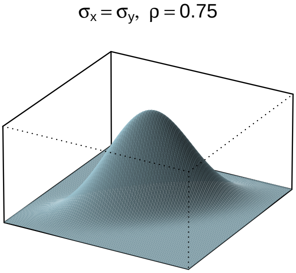
Pictures

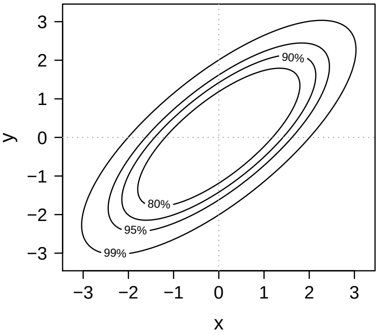
$2\sigma_x = \sigma_y, \ \rho = 0$



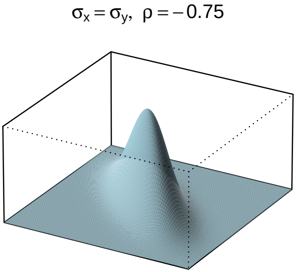


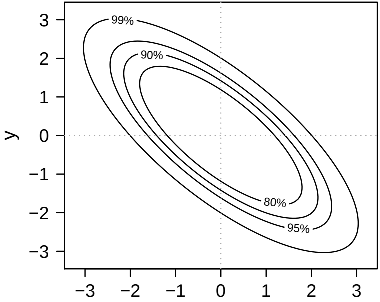
Pictures



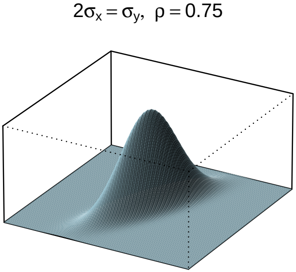


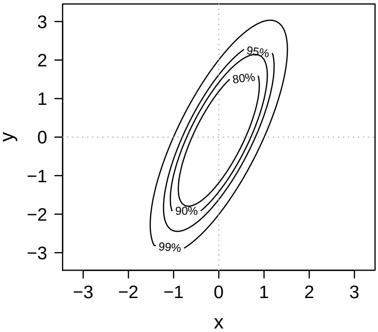
Pictures





Pictures





Comments

1. $Q(x, y) \geq 0$ with equality only when $\underline{x} = \underline{\mu}$. It follows that the density function has its mode at $\underline{x} = \underline{\mu}$.
2. Changing the values of μ_x, μ_y does not change the shape of the plots, but corresponds to a translation of the xy-plane i.e. changing μ_x, μ_y just shifts the contours / surface to a new mode position.
3. The contours of equal density are circular when $\sigma_x = \sigma_y$ and $\rho = 0$ and elliptical when $\sigma_x \neq \sigma_y$ or $\rho \neq 0$.
4. σ_x and σ_y control the extent to which the distribution is dispersed.
5. The parameter ρ is the correlation of X, Y i.e. $\mathrm{Cor}(X, Y) = \rho$. Thus for non-zero ρ, the contours are at an angle to the axes.

Marginals and conditionals

Suppose $(X, Y)^T \sim N_2(\underline{\mu}, \Sigma)$. Then:-

1. The marginal distributions are normal:

$$X \sim N(\mu_x, \sigma_x^2) \quad \text{and}$$
$$Y \sim N(\mu_y, \sigma_y^2).$$

2. The conditional distributions are normal:

$$X|Y = y \;\sim\; N(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y), \sigma_x^2(1 - \rho^2)) \quad \text{and}$$
$$Y|X = x \;\sim\; N(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1 - \rho^2)).$$

3. When $\rho = 0$, X and Y are independent.
4. Linear combinations of X and Y are also normally distributed:

$$aX + bY \sim N(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\rho\sigma_x\sigma_y)$$

where a, b are constants.

Example 4.3

Suppose $(X, Y)^T \sim N_2(\underline{\mu}, \Sigma)$ where $\mu_x = 2$, $\mu_y = 3$, $\sigma_x = 1$, $\sigma_y = 1$ and $\rho = 0.5$.

Simulate a sample of size 500 from this distribution and draw a scatter plot.

Use simulation to find $\Pr\big(X^2 + Y^2 < 9\big)$.

Solution

The marginal distribution of X is $X \sim N(2, 1^2)$.

Using the formula for the conditional

$$\begin{aligned} Y|X = x \; &\sim \; N(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1-\rho^2)) \\ &\sim \; N(3 + 0.5(x-2), 0.75). \end{aligned}$$

Example 4.3

Suppose $(X, Y)^T \sim N_2(\underline{\mu}, \Sigma)$ where $\mu_x = 2$, $\mu_y = 3$, $\sigma_x = 1$, $\sigma_y = 1$ and $\rho = 0.5$.

Simulate a sample of size 500 from this distribution and draw a scatter plot.

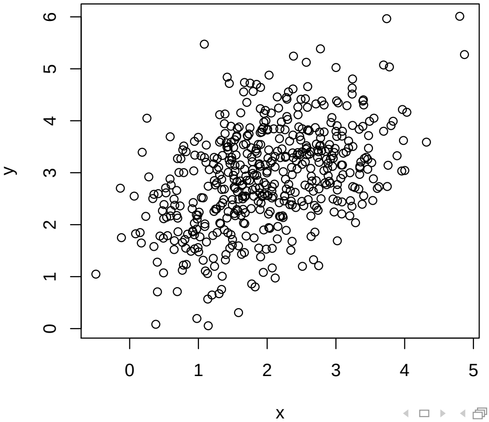
Use simulation to find $\Pr\big(X^2 + Y^2 < 9\big)$.

Solution
The marginal distribution of X is $X \sim N(2, 1^2)$.
Using the formula for the conditional

$$\begin{aligned} Y|X = x \ &\sim \ N(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1 - \rho^2)) \\ &\sim \ N(3 + 0.5(x - 2), 0.75). \end{aligned}$$

Simulation results

```
npts = 500
x = rnorm(npts, mean=2, sd = 1)
y = rnorm(npts, mean=3+0.5*(x-2), sd=sqrt(0.75))
```



Probability calculation

To find $\Pr(X^2 + Y^2 < 9)$ approximately, count the number of points in the region:

```
npts = 10000
x = rnorm(npts, mean=2, sd = 1)
y = rnorm(npts, mean=3+0.5*(x-2),sd=sqrt(0.75))
f = x^2+y^2
sum(f<9)/npts
```

Answer $\simeq$ 0.2776

The multivariate normal distribution

The multivariate normal distribution is defined on vectors in $\mathbb{R}^n$. Suppose that $\underline{X}$ is a random vector with n entries, i.e. $\underline{X} = (X_1, \ldots, X_n)^T$.

Then

$$\underline{X} \sim N_n(\underline{\mu}, \Sigma)$$

if $X_1, \ldots, X_n$ have joint PDF given by

$$f_{\underline{X}}(\underline{x}) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} Q(\underline{x})\right)$$

where

$$Q(\underline{x}) = (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu}).$$

This definition makes sense for any column vector $\underline{\mu} \in \mathbb{R}^n$ and any positive definite $n \times n$ matrix Σ.

Remarks

1. The vector $\underline{\mu}$ is the mean of the distribution and Σ is called the covariance matrix.
2. All the marginal distributions of $\underline{X}$ are normal. (We do not specify their parameters here, however).
3. Similarly, all the conditional distributions of $\underline{X}$ are normal. (Again, we do not specify the parameters of these distributions here).

Extra example

Suppose

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left[\begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 8 & 2 \\ 2 & 5 \end{pmatrix}\right].$$

The random variable Z is defined by $Z = X + 3Y$. What is the distribution of Z?

Extra example

We have $Z = X + 3Y$. Using result 4 on page 25, we have

$$E[Z] = 1 \times \mu_x + 3 \times \mu_y = 1 \times 4 + 3 \times 1 = 7.$$

Now from the variance-covariance matrix, we have $\rho\sigma_x\sigma_y = 2$. Thus

$$\begin{aligned} Var(Z) &= 1^2 \times \sigma_x^2 + 3^2 \times \sigma_y^2 + 2 \times 1 \times 3 \times (\rho\sigma_x\sigma_y) \\ &= 1 \times 8 + 9 \times 5 + 2 \times 1 \times 3 \times 2 \\ &= 65. \end{aligned}$$

Therefore $Z \sim N(7, 65)$.

Extra example

We have $Z = X + 3Y$. Using result 4 on page 25, we have

$$E[Z] = 1 \times \mu_x + 3 \times \mu_y = 1 \times 4 + 3 \times 1 = 7.$$

Now from the variance-covariance matrix, we have $\rho\sigma_x\sigma_y = 2$. Thus

$$\begin{aligned} Var(Z) &= 1^2 \times \sigma_x^2 + 3^2 \times \sigma_y^2 + 2 \times 1 \times 3 \times (\rho\sigma_x\sigma_y) \\ &= 1 \times 8 + 9 \times 5 + 2 \times 1 \times 3 \times 2 \\ &= 65. \end{aligned}$$

Therefore $Z \sim N(7, 65)$.

Extra example

We have $Z = X + 3Y$. Using result 4 on page 25, we have

$$E[Z] = 1 \times \mu_x + 3 \times \mu_y = 1 \times 4 + 3 \times 1 = 7.$$

Now from the variance-covariance matrix, we have $\rho\sigma_x\sigma_y = 2$. Thus

$$\begin{aligned} Var(Z) &= 1^2 \times \sigma_x^2 + 3^2 \times \sigma_y^2 + 2 \times 1 \times 3 \times (\rho\sigma_x\sigma_y) \\ &= 1 \times 8 + 9 \times 5 + 2 \times 1 \times 3 \times 2 \\ &= 65. \end{aligned}$$

Therefore $Z \sim N(7, 65)$.

Extra example

We have $Z = X + 3Y$. Using result 4 on page 25, we have

$$E[Z] = 1 \times \mu_x + 3 \times \mu_y = 1 \times 4 + 3 \times 1 = 7.$$

Now from the variance-covariance matrix, we have $\rho\sigma_x\sigma_y = 2$. Thus

$$\begin{aligned} Var(Z) &= 1^2 \times \sigma_x^2 + 3^2 \times \sigma_y^2 + 2 \times 1 \times 3 \times (\rho\sigma_x\sigma_y) \\ &= 1 \times 8 + 9 \times 5 + 2 \times 1 \times 3 \times 2 \\ &= 65. \end{aligned}$$

Therefore $Z \sim N(7, 65)$.

Extra example