

MAS2602: Computing for Statistics

Newcastle University
`lee.fawcett@ncl.ac.uk`

Semester 1, 2019/20

MAS2602 Arrangements

- Statistics classes run in teaching weeks 5–9
- Lecturer is Dr. Lee Fawcett (lee.fawcett@ncl.ac.uk)
- Schedule: 4 lectures, 4 practicals, 1 revision class, 1 class test
- Office hours: Tuesdays 3-4; room 2.07 Herschel
(also MAS3902: Tuesdays 2-3 and Thursdays 3-4)

Schedule

Week 5	Mon 28 Oct	11–12	Lecture (Herschel LT1)
	Fri 1 Nov	10–12	Practical (Herschel PC)
Week 6	Mon 4 Nov	11–12	Lecture (Herschel LT1)
	Wed 6 Nov	11–1	Practical (Herschel PC)
Week 7	Mon 11 Nov	11–12	Lecture (Herschel LT1)
	Thu 14 Nov	11–1	Practical (Herschel PC)
Week 8	Mon 18 Nov	11–12	Lecture (Herschel LT1)
	Thu 21 Nov	11–1	Practical (Herschel PC)
	<i>Thu 21 Nov</i>	<i>3pm</i>	<i>Assignment due</i>
	Fri 22 Nov	1–2	Revision (Herschel LT2)
Week 9	<i>Tue 26 Nov</i>	<i>9–10</i>	<i>Test (Herschel PC)</i>

Assignment (a.k.a. “mini-project”)

- Due **Thursday 21 November, 3pm**
- Worth 10% of the module marks

Class test

- **Tuesday 26 November, 9am** – one hour long
- Worth 30% of module marks
- Open-book test
- You will write one or two short R programs during the test

Help and Support

Help available:

- Office hours
- Demonstrators in practical sessions
- Books – see recommendations in booklet
- Email Lee, or just pop in!
- Blackboard and dedicated webpage

Late work policy

In this module, deadline extensions can be requested for the final project (by means of submitting a PEC form), and work submitted within 7 days of the deadline without good reason will be marked for reduced credit. This module also contains tests worth more than 10% for which rescheduling can be requested (by means of submitting a PEC form). There are mini-projects (worth 10% each) for which it is not possible to extend deadlines and for which no late work can be accepted. For details of the policy (including procedures in the event of illness etc.) please look at the School web site:

<http://www.ncl.ac.uk/maths/students/teaching/homework/>

For problems with deadlines, speak to your personal tutor and prepare a PEC form

Lecture 1: Introduction and Simulation of Random Variables

Introduction

In this part of the module we will do statistics with R:

- R is the foremost tool in modern computational statistics
- Using R teaches general concepts in programming
- It can be used to illustrate mathematical ideas in probability and statistics

Today's lecture: simulating random variables

- 1 Simulating random variables seen in MAS1604
- 2 Using this to simulate more complicated probability models

Friday's practical:

- 1 Revision of R from MAS1802
- 2 Putting today's material into practice

Introduction

In this part of the module we will do statistics with R:

- R is the foremost tool in modern computational statistics
- Using R teaches general concepts in programming
- It can be used to illustrate mathematical ideas in probability and statistics

Today's lecture: simulating random variables

- 1 Simulating random variables seen in MAS1604
- 2 Using this to simulate more complicated probability models

Friday's practical:

- 1 Revision of R from MAS1802
- 2 Putting today's material into practice

Introduction

In this part of the module we will do statistics with R:

- R is the foremost tool in modern computational statistics
- Using R teaches general concepts in programming
- It can be used to illustrate mathematical ideas in probability and statistics

Today's lecture: simulating random variables

- 1 Simulating random variables seen in MAS1604
- 2 Using this to simulate more complicated probability models

Friday's practical:

- 1 Revision of R from MAS1802
- 2 Putting today's material into practice

The binomial distribution

If $X \sim \text{Bin}(n, p)$ then X has PMF (probability mass function) given by

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k = 0, 1, \dots, n.$$

There is no closed formula for the CDF (cumulative distribution function).

R commands:

- `dbinom` – calculate PMF
- `pbinom` – calculate CDF
- `rbinom` – generate random sample

The binomial distribution

If $X \sim \text{Bin}(n, p)$ then X has PMF (probability mass function) given by

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k = 0, 1, \dots, n.$$

There is no closed formula for the CDF (cumulative distribution function).

R commands:

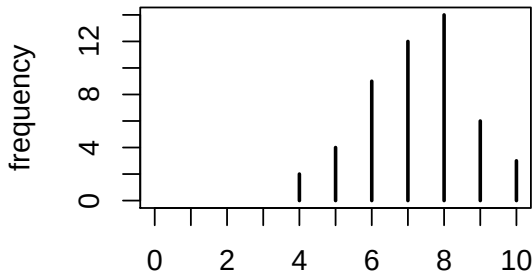
- `dbinom` – calculate PMF
- `pbinom` – calculate CDF
- `rbinom` – generate random sample

R commands for the binomial distribution

```
1 dbinom(5, 10, 0.7)      #  $\Pr(X = 5), X \sim \text{Bin}(10, 0.7)$   
2 pbinom(4, 10, 0.7)      #  $\Pr(X \leq 4), X \sim \text{Bin}(10, 0.7)$   
3 rbinom(50, 10, 0.7)     # Sample a  $\text{Bin}(10, 0.7)$  distribution 50 times
```

Creating a bar plot

```
1 x = rbinom(50, 10, 0.7)
2 plot(table(x), xlim = c(0,10), xaxt = 'n', xlab = 'x',
3      ylab = 'frequency')
4 axis(1, at = seq(0, 10))
```



The geometric distribution

If $Y \sim \text{Geom}(p)$ then Y has PMF and CDF given by

$$p_Y(k) = (1 - p)^{k-1}p,$$

$$F_Y(k) = 1 - (1 - p)^k, \quad \text{for } k = 1, 2, \dots$$

Note that Y takes values in $1, 2, 3, \dots$:

- R uses a slightly different definition
- We use the definition that Y is the number of Bernoulli trials with up to and including first success
- R counts number of trials up to, but **not** including the first success, so in R geometric random variables take values $0, 1, 2, \dots$

R commands for the geometric distribution

Adjust the arguments to account for different definition:

```
1 dgeom(4, 0.2)           #  $\Pr(Y = 5), Y \sim \text{Geom}(0.2)$   
2 pgeom(2, 0.2)           #  $\Pr(Y \leq 3), Y \sim \text{Geom}(0.2)$   
3 1 + rgeom(100, 0.2)     # Sample a  $\text{Geom}(0.2)$  distribution 100 times
```

Here's a function to replace `dgeom` with our definition of the geometric distribution:

```
1 mydgeom = function(x, p) {  
2   dgeom(x-1, p)}  
}
```

The Poisson distribution

If $Z \sim \text{Po}(\lambda)$ then Z has PMF given by

$$p_Z(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{for } k = 0, 1, 2, \dots$$

There is no closed formula for the CDF (cumulative distribution function).

R commands:

- `dpois` – calculate PMF
- `ppois` – calculate CDF
- `rpois` – generate random sample

R commands for the Poisson distribution

```
1 dpois(5, 3.5)      #  $\Pr(Z = 5), Z \sim Po(3.5)$   
2 ppois(2, 3.5)      #  $\Pr(Z \leq 2), Z \sim Po(3.5)$   
3 rpois(100, 3.5)    # Sample a  $Po(3.5)$  distribution 100 times
```

Summary

Distribution	Binomial	Poisson	Geometric 
PMF	<code>dbinom(...)</code>	<code>dpois(...)</code>	<code>dgeom(...)</code>
CDF	<code>pbinom(...)</code>	<code>ppois(...)</code>	<code>pgeom(...)</code>
sample	<code>rbinom(...)</code>	<code>rpois(...)</code>	<code>rgeom(...)</code>

Continuous random variables

R has functions for the uniform, exponential and normal distributions:

Distribution	Uniform	Exponential	Normal
PDF	<code>dunif(...)</code>	<code>dexp(...)</code>	<code>dnorm(...)</code>
CDF	<code>punif(...)</code>	<code>pexp(...)</code>	<code>pnorm(...)</code>
quantile	<code>qunif(...)</code>	<code>qexp(...)</code>	<code>qnorm(...)</code>
sample	<code>runif(...)</code>	<code>rexp(...)</code>	<code>rnorm(...)</code>

Continuous random variables – R examples

```
1 dunif(3, 2, 5)      #  $f_X(3), X \sim U(2, 5)$   
2 pexp(1.5, 5)        #  $F_Y(1.5), Y \sim \text{Exp}(5)$   
3 rnorm(10, 0, 2)     # Sample a  $N(0, 2^2)$  distribution 10 times
```

For the standard uniform ($U(0, 1)$) and standard normal ($N(0, 1)$) distributions you don't need to provide the parameters $a = 0$, $b = 1$ and $\mu = 0$, $\sigma = 1$ respectively. For example:

```
1 runif(20)           # Samples a  $U(0, 1)$  distribution 20 times  
2 pnorm(1.96)         #  $\Pr(Z < 1.96), Z \sim N(0, 1)$   
3 [1] 0.9750021
```

Quantiles

The quantile functions `qunif`, `qexp`, `qnorm` solve equations like

$$F_X(\alpha) = p$$

for α given a probability p . For example:

```
1 qnorm(0.9750021)
2 [1] 1.96
```

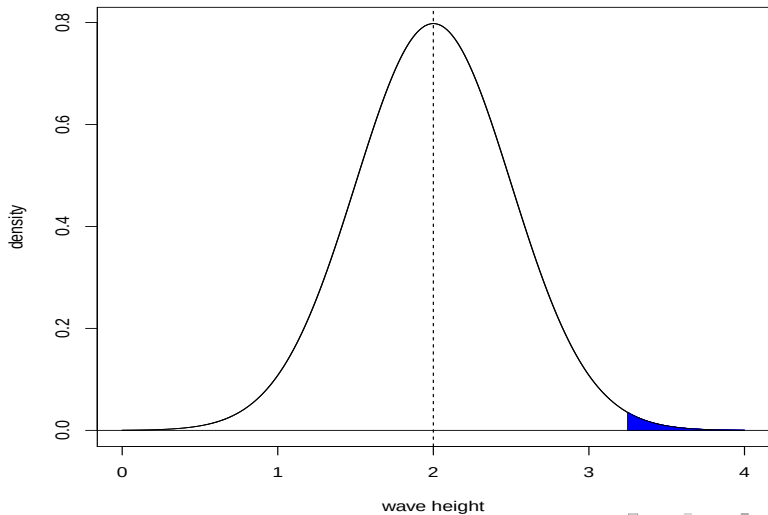
Quantiles

Example

Suppose annual maximum wave heights observed off the coast at a flood-prone town are assumed Normally distributed, with mean 2 metres and standard deviation 0.5 metres.

- (a) Write down the R command to find the probability that the largest wave height next year will exceed 3.25 metres.
- (b) Write down the R command to estimate the height of a new sea wall such that we might expect the town to be flooded, on average, once per century. Why might our modelling assumption be invalid?

Example 1.1: Solution (a)



Example 1.1: Solution (a)

Could work out the old-fashioned way:

$$\begin{aligned}\Pr(X > 3.25) &= \Pr\left(Z > \frac{3.25 - 2}{0.5}\right) \\ &= \Pr(Z > 2.5) \\ &= 1 - \Pr(Z \leq 2.5) = 1 - 0.994 = 0.006.\end{aligned}$$

Or could just use R:

```
1 1-pnorm(3.25, 2, 0.5)
2 [1] 0.006209665
```

Example 1.1: Solution (a)

Could work out the old-fashioned way:

$$\begin{aligned}\Pr(X > 3.25) &= \Pr\left(Z > \frac{3.25 - 2}{0.5}\right) \\ &= \Pr(Z > 2.5) \\ &= 1 - \Pr(Z \leq 2.5) = 1 - 0.994 = 0.006.\end{aligned}$$

Or could just use R:

```
1 1-pnorm(3.25, 2, 0.5)
2 [1] 0.006209665
```

Example 1.1: Solution (a)

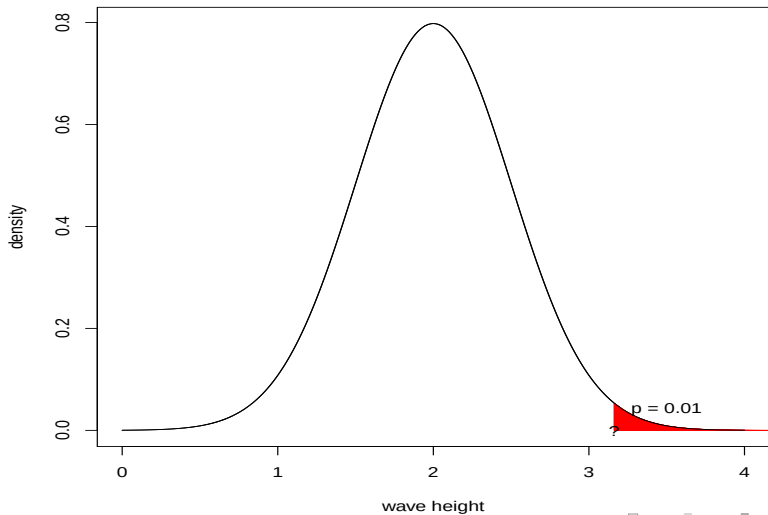
Could work out the old-fashioned way:

$$\begin{aligned}\Pr(X > 3.25) &= \Pr\left(Z > \frac{3.25 - 2}{0.5}\right) \\ &= \Pr(Z > 2.5) \\ &= 1 - \Pr(Z \leq 2.5) = 1 - 0.994 = 0.006.\end{aligned}$$

Or could just use R:

```
1 1-pnorm(3.25, 2, 0.5)
2 [1] 0.006209665
```

Example 1.1: Solution (b)



Example 1.1: Solution (b)

Similarly:

```
1 qnorm(0.99, 2, 0.5)
2 [1] 3.163174
```

A more advanced model

- Number of arrivals per day at an IT help-desk is modelled using a Poisson distribution
- Mean of the Poisson distribution might vary from day-to-day
- Suppose the number of arrivals $X \sim Po(\Lambda)$
- Λ is itself a random variable, with $\Lambda \sim Exp(c)$ for a constant $c = 0.05$

What can we say about the distribution of X ?

- 1 What are the expectation and variance of X ?
- 2 What is $\Pr(X > 30)$?

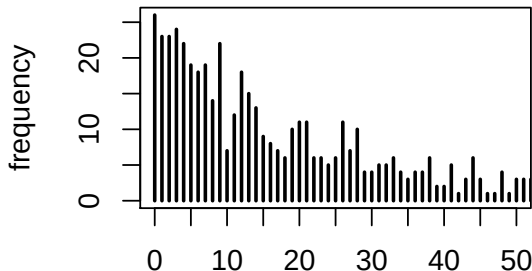
Arrival model – R code

```
1 arrivals = function(n, c = 0.05) {  
2   x = vector(mode = 'numeric', length = n)  
3   for (i in 1:n) {  
4     lambda = rexp(1, rate = c)  
5     x[i] = rpois(1, lambda)  
6   }  
7   return(x)  
8 }
```

Arrival model – R code

Bar plot:

```
1 x = arrivals(500)
2 plot(table(x), xlim = c(0, 50), xaxt = 'n', xlab = 'x'
3     , ylab = 'frequency')
4 axis(1, at = seq(0, 50, 5))
```



Using simulated samples

- The **sample mean** is an approximation to the distribution mean
- The same applies for the sample variance
- To calculate $\Pr(X > 30)$ we count the proportion of times this occurs in the sample

We expect the approximation to improve **as we increase the sample size**

Using simulated samples

```
1 mean(x)
2 [1] 19.618
3 var(x)
4 [1] 382.1764
5 sum(x>30)/500
6 [1] 0.202
```

Mixtures of normal distributions

Suppose

- $\mu_1, \mu_2, \sigma_1 > 0, \sigma_2 > 0$ are fixed constants,
- w_1, w_2 are positive constants with $w_1 + w_2 = 1$.

Consider the following function:

$$f(x) = w_1 f_1(x) + w_2 f_2(x)$$

where f_1 and f_2 are the density functions for $Z_1 \sim N(\mu_1, \sigma_1^2)$ and $Z_2 \sim N(\mu_2, \sigma_2^2)$ respectively.

Check that $f(x)$ represents a valid probability density function.

Mixtures of normal distributions

First note that $f(x) \geq 0$ everywhere. Also

$$\begin{aligned}\int_{-\infty}^{\infty} f(x) dx &= w_1 \int_{-\infty}^{\infty} f_1(x) dx + w_2 \int_{-\infty}^{\infty} f_2(x) dx \\ &= w_1 + w_2 = 1.\end{aligned}$$

Mixtures of normal distributions

First note that $f(x) \geq 0$ everywhere. Also

$$\begin{aligned}\int_{-\infty}^{\infty} f(x) dx &= w_1 \int_{-\infty}^{\infty} f_1(x) dx + w_2 \int_{-\infty}^{\infty} f_2(x) dx \\ &= w_1 + w_2 = 1.\end{aligned}$$

Mixtures of normal distributions

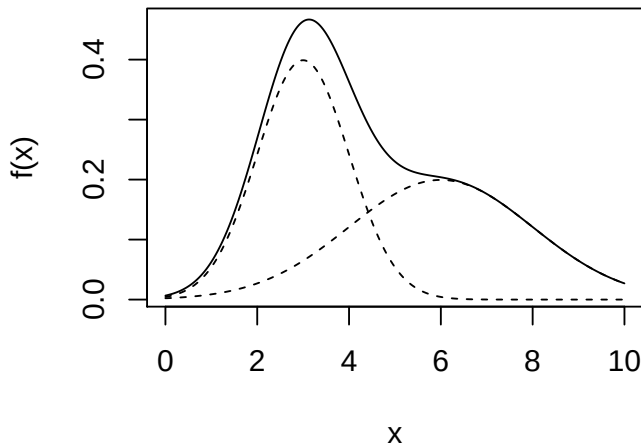
If a random variable X has PDF corresponding to $f(x)$ we say it is a **mixture** of normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ with weights w_1, w_2 .

Note that X is **not** the sum of two normal random variables i.e.

$$X \neq w_1 Z_1 + w_2 Z_2 \quad \text{where } Z_i \sim N(\mu_i, \sigma_i^2), \quad i = 1, 2.$$

Example

For example: suppose $\mu_1 = 3, \sigma_1 = 1$ and $\mu_2 = 6, \sigma_2 = 2$ with $w_1 = w_2 = 1/2$.



Sampling from a mixture distribution

- 1 Sample a random variable $J \in \{1, 2\}$ such that $\Pr(J = 1) = w_1$ and $\Pr(J = 2) = w_2$.
- 2 The random variable J tells you which **component** of the mixture to sample X from.
- 3 If $J = 1$ then sample X from $N(\mu_1, \sigma_1^2)$, but if $J = 2$ then sample X from $N(\mu_2, \sigma_2^2)$.

Sampling from a mixture distribution

```
1  normal.mixture = function(n, mu1, sig1, w1, mu2, sig2,
2    w2) {
3    p = c(w1, w2)
4    x = vector(mode = 'numeric', length = n)
5    for (i in 1:n) {
6      j = sample(c(1, 2), 1, prob = p)
7      if (j==1) {
8        x[i] = rnorm(1, mu1, sig1)
9      }
10     else {
11       x[i] = rnorm(1, mu2, sig2)
12     }
13   }
14   return(x)
15 }
```

Lecture 2: Inverse CDF method

Today's lecture

In this lecture we look at the **inverse CDF method** for simulating from continuous random variables.

The inverse CDF method

This is a method for simulating **univariate continuous** random variables

- Let $U \sim U(0, 1)$
- Suppose $F(x)$ is a well-defined CDF which is invertible
- Then the random variable $X = F^{-1}(U)$ has CDF $F(x)$

The inverse CDF method

$$X = F^{-1}(U).$$

So

$$\begin{aligned}\Pr(X \leq x) &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(F(F^{-1}(U)) \leq F(x)) \\ &= \Pr(U \leq F(x)) \\ &= F(x),\end{aligned}$$

since $0 \leq F(x) \leq 1$ (see sketch on board).

The inverse CDF method

$$X = F^{-1}(U).$$

So

$$\begin{aligned}\Pr(X \leq x) &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(F(F^{-1}(U)) \leq F(x)) \\ &= \Pr(U \leq F(x)) \\ &= F(x),\end{aligned}$$

since $0 \leq F(x) \leq 1$ (see sketch on board).

The inverse CDF method

$$X = F^{-1}(U).$$

So

$$\begin{aligned}\Pr(X \leq x) &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(F(F^{-1}(U)) \leq F(x)) \\ &= \Pr(U \leq F(x)) \\ &= F(x),\end{aligned}$$

since $0 \leq F(x) \leq 1$ (see sketch on board).

The inverse CDF method

$$X = F^{-1}(U).$$

So

$$\begin{aligned}\Pr(X \leq x) &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(F(F^{-1}(U)) \leq F(x)) \\ &= \Pr(U \leq F(x)) \\ &= F(x),\end{aligned}$$

since $0 \leq F(x) \leq 1$ (see sketch on board).

The inverse CDF method

$$X = F^{-1}(U).$$

So

$$\begin{aligned}\Pr(X \leq x) &= \Pr(F^{-1}(U) \leq x) \\ &= \Pr(F(F^{-1}(U)) \leq F(x)) \\ &= \Pr(U \leq F(x)) \\ &= F(x),\end{aligned}$$

since $0 \leq F(x) \leq 1$ (see sketch on board).

The inverse CDF method

Example

Suppose we obtain two observations from a $U(0, 1)$ distribution: 0.1 and 0.85. Use these values to obtain two observations from $X \sim \text{Exp}(2)$.

Solution

If $X \sim \text{Exp}(2)$ then $F_X(x) = 1 - e^{-2x}$.

If $y = 1 - e^{-2x}$, then

$$e^{-2x} = 1 - y$$

$$-2x = \log(1 - y)$$

$$x = -\frac{1}{2} \log(1 - y).$$

Plug in $y = 0.1$ and $y = 0.85$; gives realized values of 0.0527 and 0.949.

Solution

If $X \sim \text{Exp}(2)$ then $F_X(x) = 1 - e^{-2x}$.

If $y = 1 - e^{-2x}$, then

$$e^{-2x} = 1 - y$$

$$-2x = \log(1 - y)$$

$$x = -\frac{1}{2} \log(1 - y).$$

Plug in $y = 0.1$ and $y = 0.85$; gives realized values of 0.0527 and 0.949.

Solution

If $X \sim \text{Exp}(2)$ then $F_X(x) = 1 - e^{-2x}$.

If $y = 1 - e^{-2x}$, then

$$e^{-2x} = 1 - y$$

$$-2x = \log(1 - y)$$

$$x = -\frac{1}{2} \log(1 - y).$$

Plug in $y = 0.1$ and $y = 0.85$; gives realized values of 0.0527 and 0.949.

Example

- Patients arriving at a doctor's surgery have appointment times 15 minutes apart.
- The random variable L measures how late a patient is for his/her appointment, with negative values denoting early arrivals.

The PDF of L is

$$f_L(\ell) = \begin{cases} 0.0025(\ell + 20), & -20 \leq \ell \leq 0, \\ 0.05e^{-\ell/10}, & \ell > 0. \end{cases}$$

Questions

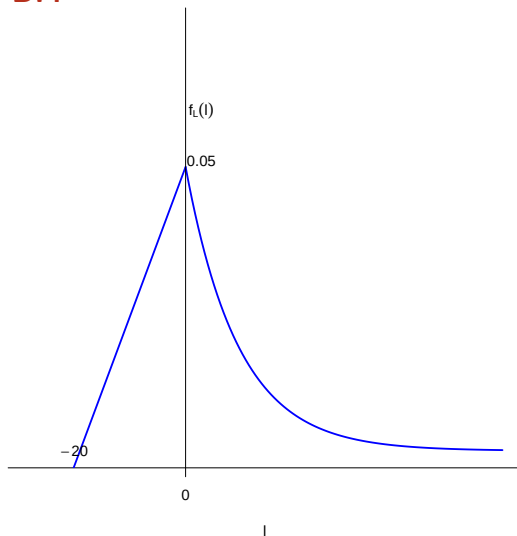
- a. Sketch $f_L(\ell)$.
- b. Show that the CDF of L is

$$F_L(\ell) = \begin{cases} 0, & \ell < -20, \\ 0.00125(\ell + 20)^2, & -20 \leq \ell \leq 0, \\ 1 - \frac{e^{-\ell/10}}{2}, & \ell > 0. \end{cases}$$

- c. Suppose that 0.205 and 0.713 are two realized values of $U \sim U(0, 1)$. Show that the corresponding realized values of L are -7.19 and 5.55 under the inverse CDF simulation method.

Solution

First, the PDF:



Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now to find the CDF: L lies between -20 and $+\infty$, so $F_L(\ell) = 0$ when $\ell < -20$.

For $-20 \leq \ell \leq 0$:

$$\begin{aligned} F_L(\ell) &= \int_{-20}^{\ell} f_L(s) ds \\ &= \int_{-20}^{\ell} 0.0025(s + 20) ds \\ &= \left[\frac{0.0025}{2} (s + 20)^2 \right]_{-20}^{\ell} \\ &= 0.00125(\ell + 20)^2 \quad \text{for } -20 \leq \ell \leq 0. \end{aligned}$$

Also, $F_L(0) = 0.5$.

Solution

Now

$$\begin{aligned}F_L(\ell) &= \int_{-20}^0 f_L(s)ds + \int_0^\ell f_L(s)ds \quad \text{for } \ell > 0 \\&= 0.5 + \int_0^\ell 0.05e^{-s/10}ds \\&= 0.5 + \left[0.05 \times -10e^{-s/10}\right]_0^\ell \\&= 0.5 + 0.5 - 0.5e^{-\ell/10} \\&= 1 - \frac{e^{-\ell/10}}{2}, \quad \text{as required.}\end{aligned}$$

Solution

Now

$$\begin{aligned}F_L(\ell) &= \int_{-20}^0 f_L(s)ds + \int_0^{\ell} f_L(s)ds \quad \text{for } \ell > 0 \\&= 0.5 + \int_0^{\ell} 0.05e^{-s/10}ds \\&= 0.5 + \left[0.05 \times -10e^{-s/10}\right]_0^{\ell} \\&= 0.5 + 0.5 - 0.5e^{-\ell/10} \\&= 1 - \frac{e^{-\ell/10}}{2}, \quad \text{as required.}\end{aligned}$$

Solution

Now

$$\begin{aligned}F_L(\ell) &= \int_{-20}^0 f_L(s)ds + \int_0^{\ell} f_L(s)ds \quad \text{for } \ell > 0 \\&= 0.5 + \int_0^{\ell} 0.05e^{-s/10}ds \\&= 0.5 + \left[0.05 \times -10e^{-s/10}\right]_0^{\ell} \\&= 0.5 + 0.5 - 0.5e^{-\ell/10} \\&= 1 - \frac{e^{-\ell/10}}{2}, \quad \text{as required.}\end{aligned}$$

Solution

Now

$$\begin{aligned}F_L(\ell) &= \int_{-20}^0 f_L(s)ds + \int_0^{\ell} f_L(s)ds \quad \text{for } \ell > 0 \\&= 0.5 + \int_0^{\ell} 0.05e^{-s/10}ds \\&= 0.5 + \left[0.05 \times -10e^{-s/10}\right]_0^{\ell} \\&= 0.5 + 0.5 - 0.5e^{-\ell/10} \\&= 1 - \frac{e^{-\ell/10}}{2}, \quad \text{as required.}\end{aligned}$$

Solution

Now

$$\begin{aligned}F_L(\ell) &= \int_{-20}^0 f_L(s)ds + \int_0^\ell f_L(s)ds \quad \text{for } \ell > 0 \\&= 0.5 + \int_0^\ell 0.05e^{-s/10}ds \\&= 0.5 + \left[0.05 \times -10e^{-s/10}\right]_0^\ell \\&= 0.5 + 0.5 - 0.5e^{-\ell/10} \\&= 1 - \frac{e^{-\ell/10}}{2}, \quad \text{as required.}\end{aligned}$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Now for some realized values:

$u = 0.205 < 0.5$, so realized value ℓ lies between -20 and 0 . Thus

$$F_L(\ell) = 0.00125(\ell + 20)^2.$$

Find the inverse, and evaluate at $u = 0.205$:

$$y = 0.00125(\ell + 20)^2, \quad \text{so}$$

$$\ell = \pm \sqrt{\frac{y}{0.00125}} - 20,$$

giving

$$\ell = \sqrt{\frac{0.205}{0.00125}} - 20 = -7.2.$$

Solution

Similarly, $u = 0.713 > 0.5$ so realized value > 0 . Thus

$$F_L(\ell) = 1 - \frac{e^{-\ell/10}}{2}.$$

Find the inverse, and evaluate at $u = 0.713$:

$$\begin{aligned} y &= 1 - \frac{e^{-\ell/10}}{2} \quad \text{so} \\ \ell &= -10 \log(2 - 2y), \end{aligned}$$

giving

$$\ell = -10 \log(2 - 2 \times 0.713) = 5.55.$$

Solution

Similarly, $u = 0.713 > 0.5$ so realized value > 0 . Thus

$$F_L(\ell) = 1 - \frac{e^{-\ell/10}}{2}.$$

Find the inverse, and evaluate at $u = 0.713$:

$$\begin{aligned} y &= 1 - \frac{e^{-\ell/10}}{2} \quad \text{so} \\ \ell &= -10 \log(2 - 2y), \end{aligned}$$

giving

$$\ell = -10 \log(2 - 2 \times 0.713) = 5.55.$$

Solution

Similarly, $u = 0.713 > 0.5$ so realized value > 0 . Thus

$$F_L(\ell) = 1 - \frac{e^{-\ell/10}}{2}.$$

Find the inverse, and evaluate at $u = 0.713$:

$$\begin{aligned} y &= 1 - \frac{e^{-\ell/10}}{2} \quad \text{so} \\ \ell &= -10 \log(2 - 2y), \end{aligned}$$

giving

$$\ell = -10 \log(2 - 2 \times 0.713) = 5.55.$$

Solution

Similarly, $u = 0.713 > 0.5$ so realized value > 0 . Thus

$$F_L(\ell) = 1 - \frac{e^{-\ell/10}}{2}.$$

Find the inverse, and evaluate at $u = 0.713$:

$$\begin{aligned} y &= 1 - \frac{e^{-\ell/10}}{2} \quad \text{so} \\ \ell &= -10 \log(2 - 2y), \end{aligned}$$

giving

$$\ell = -10 \log(2 - 2 \times 0.713) = 5.55.$$

Solution

Similarly, $u = 0.713 > 0.5$ so realized value > 0 . Thus

$$F_L(\ell) = 1 - \frac{e^{-\ell/10}}{2}.$$

Find the inverse, and evaluate at $u = 0.713$:

$$\begin{aligned} y &= 1 - \frac{e^{-\ell/10}}{2} \quad \text{so} \\ \ell &= -10 \log(2 - 2y), \end{aligned}$$

giving

$$\ell = -10 \log(2 - 2 \times 0.713) = 5.55.$$

Recap Quiz

1. Write down R code that simulates 100 observations from the $U(0, 1)$ distribution and stores the output in the vector U .
2. Suppose X has CDF

$$F_X(x) = 1 - \frac{1}{x}, \quad x \geq 1.$$

Find $F_X^{-1}(x)$.

3. Write an R function called `inv.cdf` to return the inverse of $F_X(x)$.
4. Write down a single line of R code to generate 100 realizations of the random variable X .
5. Write down a single line of R code to produce a histogram of your realizations of X , coloured yellow and with a suitable title.

Lecture 3: Transformation of Random Variables

Transforming a normal random variable

- Consider how we might model a **continuous positive** random quantity e.g. blood pressure or height
- The Normal distribution is not suitable!

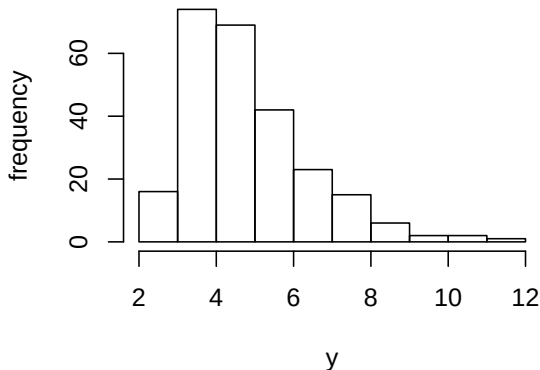
One way to obtain a positive random variable is to define

$$Y = \exp(X), \quad \text{where } X \sim N(\mu, \sigma^2).$$

Then Y is a random variable which is a **transformation** of X .

Simulation to explore transformation

```
1 y = exp(rnorm(250, mean = 1.5, sd = 0.3))  
2 hist(y, xlab = 'Y', ylab = 'frequency')
```



Calculating PDFs

- This looks like a useful distribution – what is its PDF?
- $Y = \exp(X)$ so $X = \log(Y)$
- So $f_Y(y) = f_X(\log(y))$ is a good guess, but is **NOT** the right answer!

Example

If $X \sim U(0, 1)$, what is the PDF of $Y = X^3$?

Solution

$$\begin{aligned}F_Y(y) &= \Pr(Y \leq y) \\&= \Pr(X^3 \leq y) \\&= \Pr(X \leq y^{1/3}) \\&= F_X(y^{1/3}).\end{aligned}$$

Therefore

$$\begin{aligned}f_Y(y) &= \frac{d}{dy} F_X(y^{1/3}) \\&= f_X(y^{1/3}) \times \frac{1}{3} y^{-2/3} \quad \text{by chain rule} \\&= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1 \text{ (and zero otherwise).}\end{aligned}$$

Solution

$$\begin{aligned}F_Y(y) &= \Pr(Y \leq y) \\&= \Pr(X^3 \leq y) \\&= \Pr(X \leq y^{1/3}) \\&= F_X(y^{1/3}).\end{aligned}$$

Therefore

$$\begin{aligned}f_Y(y) &= \frac{d}{dy} F_X(y^{1/3}) \\&= f_X(y^{1/3}) \times \frac{1}{3} y^{-2/3} \quad \text{by chain rule} \\&= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1 \text{ (and zero otherwise).}\end{aligned}$$

Solution

$$\begin{aligned}F_Y(y) &= \Pr(Y \leq y) \\&= \Pr(X^3 \leq y) \\&= \Pr(X \leq y^{1/3}) \\&= F_X(y^{1/3}).\end{aligned}$$

Therefore

$$\begin{aligned}f_Y(y) &= \frac{d}{dy} F_X(y^{1/3}) \\&= f_X(y^{1/3}) \times \frac{1}{3} y^{-2/3} \quad \text{by chain rule} \\&= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1 \text{ (and zero otherwise).}\end{aligned}$$

Solution

$$\begin{aligned}F_Y(y) &= \Pr(Y \leq y) \\&= \Pr(X^3 \leq y) \\&= \Pr(X \leq y^{1/3}) \\&= F_X(y^{1/3}).\end{aligned}$$

Therefore

$$\begin{aligned}f_Y(y) &= \frac{d}{dy} F_X(y^{1/3}) \\&= f_X(y^{1/3}) \times \frac{1}{3} y^{-2/3} \quad \text{by chain rule} \\&= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1 \text{ (and zero otherwise).}\end{aligned}$$

Solution

$$\begin{aligned}F_Y(y) &= \Pr(Y \leq y) \\&= \Pr(X^3 \leq y) \\&= \Pr(X \leq y^{1/3}) \\&= F_X(y^{1/3}).\end{aligned}$$

Therefore

$$\begin{aligned}f_Y(y) &= \frac{d}{dy} F_X(y^{1/3}) \\&= f_X(y^{1/3}) \times \frac{1}{3} y^{-2/3} \quad \text{by chain rule} \\&= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1 \text{ (and zero otherwise).}\end{aligned}$$

Solution

$$\begin{aligned}F_Y(y) &= \Pr(Y \leq y) \\&= \Pr(X^3 \leq y) \\&= \Pr(X \leq y^{1/3}) \\&= F_X(y^{1/3}).\end{aligned}$$

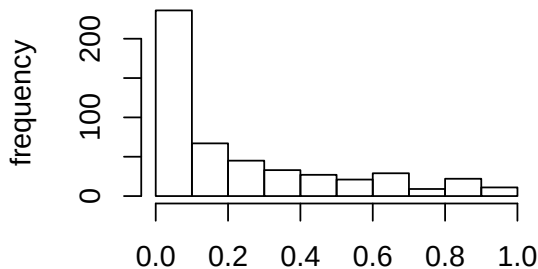
Therefore

$$\begin{aligned}f_Y(y) &= \frac{d}{dy} F_X(y^{1/3}) \\&= f_X(y^{1/3}) \times \frac{1}{3} y^{-2/3} \quad \text{by chain rule} \\&= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1 \text{ (and zero otherwise).}\end{aligned}$$

Checking via simulation

To confirm this result, we can simulate Y :

```
1 y = (runif(500))^3
2 hist(y, xlab = 'Y', ylab = 'frequency')
```



Further examples

Example

If $X \sim N(0, 1)$, what is the PDF of $Y = X^2$?

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^{1/2}) = F_X(y^{1/2}).$$

Then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(y^{1/2}) \\ &= f_X(y^{1/2}) \times \frac{1}{2} y^{-1/2} \quad \text{by chain rule} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2} y^{-1/2} \\ &= \frac{1}{2\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y \geq 0, \text{ zero otherwise.} \end{aligned}$$

Further examples

Example

If $X \sim N(0, 1)$, what is the PDF of $Y = X^2$?

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^{1/2}) = F_X(y^{1/2}).$$

Then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(y^{1/2}) \\ &= f_X(y^{1/2}) \times \frac{1}{2} y^{-1/2} \quad \text{by chain rule} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2} y^{-1/2} \\ &= \frac{1}{2\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y \geq 0, \text{ zero otherwise.} \end{aligned}$$

Further examples

Example

If $X \sim N(0, 1)$, what is the PDF of $Y = X^2$?

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^{1/2}) = F_X(y^{1/2}).$$

Then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(y^{1/2}) \\ &= f_X(y^{1/2}) \times \frac{1}{2} y^{-1/2} \quad \text{by chain rule} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2} y^{-1/2} \\ &= \frac{1}{2\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y \geq 0, \text{ zero otherwise.} \end{aligned}$$

Further examples

Example

If $X \sim N(0, 1)$, what is the PDF of $Y = X^2$?

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^{1/2}) = F_X(y^{1/2}).$$

Then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(y^{1/2}) \\ &= f_X(y^{1/2}) \times \frac{1}{2} y^{-1/2} \quad \text{by chain rule} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2} y^{-1/2} \\ &= \frac{1}{2\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y \geq 0, \text{ zero otherwise.} \end{aligned}$$

Further examples

Example

If $X \sim N(0, 1)$, what is the PDF of $Y = X^2$?

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^{1/2}) = F_X(y^{1/2}).$$

Then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(y^{1/2}) \\ &= f_X(y^{1/2}) \times \frac{1}{2} y^{-1/2} \quad \text{by chain rule} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2} y^{-1/2} \\ &= \frac{1}{2\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y \geq 0, \text{ zero otherwise.} \end{aligned}$$

Further examples

Example

If $X \sim N(0, 1)$, what is the PDF of $Y = X^2$?

$$F_Y(y) = \Pr(Y \leq y) = \Pr(X \leq y^{1/2}) = F_X(y^{1/2}).$$

Then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(y^{1/2}) \\ &= f_X(y^{1/2}) \times \frac{1}{2} y^{-1/2} \quad \text{by chain rule} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \times \frac{1}{2} y^{-1/2} \\ &= \frac{1}{2\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y \geq 0, \text{ zero otherwise.} \end{aligned}$$

Further examples

Example

If $X \sim N(\mu, \sigma^2)$, what is the PDF of $Y = \exp(X)$?

This is the example used at the start of the lecture. We say Y has the **log-normal** distribution and write $Y \sim LN(\mu, \sigma^2)$.

$$F_Y(y) = F_X(\log y) \quad \text{by usual argument with CDF}$$

$$f_Y(y) = \frac{d}{dy} F_X(\log y)$$

$$= f_X(\log y) \times \frac{1}{y} \quad \text{for } y \geq 0$$

$$= \frac{1}{\sigma y \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(\log y - \mu)^2}{\sigma^2} \right) \right\}.$$

Further examples

Example

If $X \sim N(\mu, \sigma^2)$, what is the PDF of $Y = \exp(X)$?

This is the example used at the start of the lecture. We say Y has the **log-normal** distribution and write $Y \sim LN(\mu, \sigma^2)$.

$$F_Y(y) = F_X(\log y) \quad \text{by usual argument with CDF}$$

$$f_Y(y) = \frac{d}{dy} F_X(\log y)$$

$$= f_X(\log y) \times \frac{1}{y} \quad \text{for } y \geq 0$$

$$= \frac{1}{\sigma y \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(\log y - \mu)^2}{\sigma^2} \right) \right\}.$$

Further examples

Example

If $X \sim N(\mu, \sigma^2)$, what is the PDF of $Y = \exp(X)$?

This is the example used at the start of the lecture. We say Y has the **log-normal** distribution and write $Y \sim LN(\mu, \sigma^2)$.

$$F_Y(y) = F_X(\log y) \quad \text{by usual argument with CDF}$$

$$f_Y(y) = \frac{d}{dy} F_X(\log y)$$

$$= f_X(\log y) \times \frac{1}{y} \quad \text{for } y \geq 0$$

$$= \frac{1}{\sigma y \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(\log y - \mu)^2}{\sigma^2} \right) \right\}.$$

Further examples

Example

If $X \sim N(\mu, \sigma^2)$, what is the PDF of $Y = \exp(X)$?

This is the example used at the start of the lecture. We say Y has the **log-normal** distribution and write $Y \sim LN(\mu, \sigma^2)$.

$$F_Y(y) = F_X(\log y) \quad \text{by usual argument with CDF}$$

$$f_Y(y) = \frac{d}{dy} F_X(\log y)$$

$$= f_X(\log y) \times \frac{1}{y} \quad \text{for } y \geq 0$$

$$= \frac{1}{\sigma y \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(\log y - \mu)^2}{\sigma^2} \right) \right\}.$$

Further examples

Example

If $X \sim N(\mu, \sigma^2)$, what is the PDF of $Y = \exp(X)$?

This is the example used at the start of the lecture. We say Y has the **log-normal** distribution and write $Y \sim LN(\mu, \sigma^2)$.

$$\begin{aligned} F_Y(y) &= F_X(\log y) \quad \text{by usual argument with CDF} \\ f_Y(y) &= \frac{d}{dy} F_X(\log y) \\ &= f_X(\log y) \times \frac{1}{y} \quad \text{for } y \geq 0 \\ &= \frac{1}{\sigma y \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{(\log y - \mu)^2}{\sigma^2} \right) \right\}. \end{aligned}$$

The univariate transformation method

If X has pdf $f_X(x)$ and $h(x)$ is a monotonically increasing or monotonically decreasing function then

- h is invertible and
- the PDF of $Y = h(X)$ is

$$f_Y(y) = f_X(h^{-1}(y)) \times \left| \frac{d(h^{-1}(y))}{dy} \right|$$

when y is in the range of $h(x)$ and zero otherwise.

This result allows direct calculation of $f_Y(y)$ and is called the **univariate transformation method**.

See whiteboard for Proof.

Some Examples

Example

When $X \sim U(0, 1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1. \end{aligned}$$

Some Examples

Example

When $X \sim U(0, 1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1. \end{aligned}$$

Some Examples

Example

When $X \sim U(0, 1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1. \end{aligned}$$

Some Examples

Example

When $X \sim U(0, 1)$, use the transformation method to find the PDF of $Y = X^3$.

$$h(x) = x^3, \quad h^{-1}(y) = y^{1/3}.$$

$$\begin{aligned} f_Y(y) &= 1 \times \frac{d}{dy} y^{1/3} \\ &= \frac{1}{3} y^{-2/3}, \quad 0 \leq y \leq 1. \end{aligned}$$

Some examples

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y .

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Some examples

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y .

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Some examples

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y .

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Some examples

Example

When $X \sim U(-1, 1)$ and $Y = \exp(X)$, use the transformation method to find the PDF of Y .

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{2} \times \left| \frac{d}{dy} \log(y) \right| \\ &= \frac{1}{2|y|} = \frac{1}{2y}, \quad e^{-1} \leq y \leq e^1, \text{ zero otherwise.} \end{aligned}$$

Some examples

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \quad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log(y) - \mu}{\sigma} \right)^2 \right]. \end{aligned}$$

Some examples

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \quad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log(y) - \mu}{\sigma} \right)^2 \right]. \end{aligned}$$

Some examples

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \quad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log(y) - \mu}{\sigma} \right)^2 \right]. \end{aligned}$$

Some examples

Example

Use the transformation method to find the PDF of the log-normal distribution.

$$h(x) = e^x, \quad h^{-1}(y) = \log(y).$$

$$\begin{aligned} f_Y(y) &= f_X(\log(y)) \times \frac{1}{y} \quad \text{for } y > 0 \\ &= \frac{1}{y} \times \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log(y) - \mu}{\sigma} \right)^2 \right]. \end{aligned}$$

Cauchy example

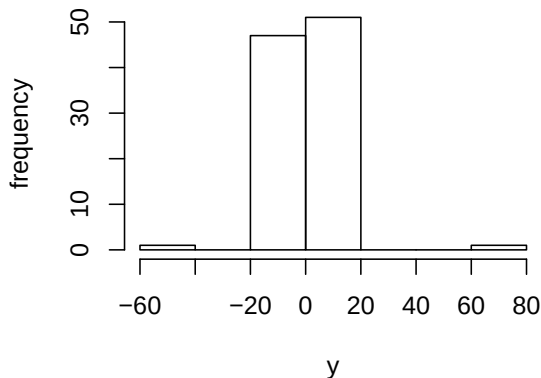
Example

Suppose $X \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$ and $Y = \tan(X)$.

Use simulation to study the distribution of Y and explicitly compute the PDF.

Solution

```
1 y = tan(runif(100, -pi/2, pi/2))  
2 hist(y, xlab = 'y', ylab = 'frequency', main = '')
```



Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi} \frac{d}{dy} (\tan^{-1} y) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi} \frac{d}{dy} (\tan^{-1} y) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has very heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi} \frac{d}{dy} (\tan^{-1} y) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi} \frac{d}{dy} (\tan^{-1} y) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Solution

This distribution looks odd! Calculate the PDF:

$$h(x) = \tan(x), \quad h^{-1}(y) = \tan^{-1}(y).$$

$$\begin{aligned} f_Y(y) &= \frac{1}{\pi} \frac{d}{dy} (\tan^{-1} y) \quad -\infty < y < \infty, \\ &= \frac{1}{\pi(1+y^2)} \quad -\infty < y < \infty. \end{aligned}$$

This is the **Cauchy distribution**. It has *very* heavy tails.

Lecture 4: Bivariate Distributions

Today's lecture

In this lecture we consider two related topics:

- 1 How to simulate samples from continuous bivariate distributions
- 2 Transformations of bivariate random variables.

This leads naturally to consideration of the bivariate normal distribution, in the next lecture.

Some other stuff first...

■ Assignment 2

- You now have the sheet (this was made available last Friday in BB)
- Due in by **4pm, Friday 30th November** (note revised deadline)
- Worth 10% of the module, so put some effort in!

■ Practical 3

- This coming Thursday, 4-6pm
- Harder than the work in previous practicals
- Based heavily on what we will cover today

Some other stuff first...

■ Assignment 2

- You now have the sheet (this was made available last Friday in BB)
- Due in by **4pm, Friday 30th November** (note revised deadline)
- Worth 10% of the module, so put some effort in!

■ Practical 3

- This coming Thursday, 4-6pm
- Harder than the work in previous practicals
- Based heavily on what we will cover today

Some other stuff first...

■ Assignment 2

- You now have the sheet (this was made available last Friday in BB)
- Due in by **4pm, Friday 30th November** (note revised deadline)
- Worth 10% of the module, so put some effort in!

■ Practical 3

- This coming Thursday, 4-6pm
- Harder than the work in previous practicals
- Based heavily on what we will cover today

Some other stuff first...

■ Assignment 2

- You now have the sheet (this was made available last Friday in BB)
- Due in by **4pm, Friday 30th November** (note revised deadline)
- Worth 10% of the module, so put some effort in!

■ Practical 3

- This coming Thursday, 4-6pm
- Harder than the work in previous practicals
- Based heavily on what we will cover today

Some other stuff first...

■ Assignment 2

- You now have the sheet (this was made available last Friday in BB)
- Due in by **4pm, Friday 30th November** (note revised deadline)
- Worth 10% of the module, so put some effort in!

■ Practical 3

- This coming Thursday, 4-6pm
- Harder than the work in previous practicals
- Based heavily on what we will cover today

Some other stuff first...

■ Assignment 2

- You now have the sheet (this was made available last Friday in BB)
- Due in by **4pm, Friday 30th November** (note revised deadline)
- Worth 10% of the module, so put some effort in!

■ Practical 3

- This coming Thursday, 4-6pm
- Harder than the work in previous practicals
- Based heavily on what we will cover today

Some other stuff first...

■ Next week

- Last week of R :-)
- Two lectures (Monday and Tuesday)
- R test on Thursday, 4-5pm - will discuss some specifics in this week's practical
- Assignment due on Friday (should have no problem with question 3 after Monday's lecture)

Some other stuff first...

■ Next week

- Last week of R :-(
- Two lectures (Monday and Tuesday)
- R test on Thursday, 4-5pm - will discuss some specifics in this week's practical
- Assignment due on Friday (should have no problem with question 3 after Monday's lecture)

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

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

where $f_X(x)$ is the **marginal** PDF of X and $f_{Y|X}(y | X = x)$ 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

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

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

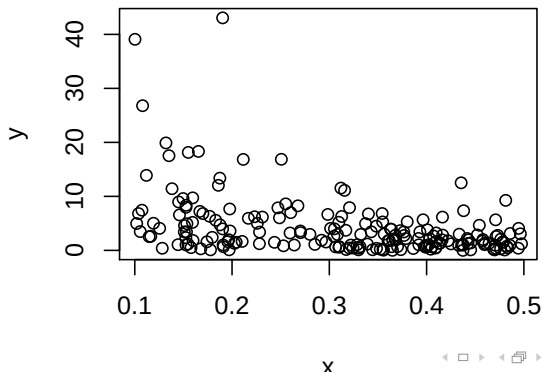
```
1 example4.1 = function(n) {  
2   output = matrix(0, 2, n)  
3   x = runif(n, 0.1, 0.5)  
4   y = rexp(n, rate = x)  
5   output[1, ] = x  
6   output[2, ] = y  
7   output  
8 }
```

Make some notes about the R function.

Example 4.1 – plot

To obtain a scatterplot:

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



The trivariate case

Example

Suppose that $X \sim U(-1, 1)$ and $Y \sim \text{Exp}(\lambda)$ 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:

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

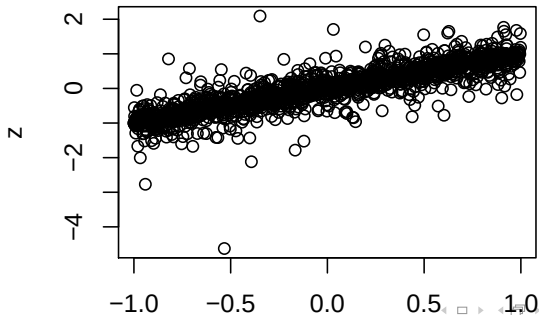
Example 4.2 – R function

```
1 example4.2 = function(n) {  
2   output = matrix(0, 3, n)  
3   x = runif(n, -1, 1)  
4   y = rexp(n, rate = 5)  
5   z = rnorm(n, mean = x, sd = y)  
6   output[1, ] = x  
7   output[2, ] = y  
8   output[3, ] = z  
9   output  
10 }
```

Make some notes about the R function.

Example 4.2 – scatterplot

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



Motivating example: points on a disc

- Consider distributions of points that lie on a disc of radius 1.
- Suppose (R, Θ) are the polar coordinates of a random point.
- Suppose that $R \sim U(0, 1)$ and $\Theta \sim U(0, 2\pi)$.

What does this distribution look like? Is it uniform over the disc?

In other words, R and Θ are uniformly distributed... are X and Y ?

Motivating example: points on a disc

- Consider distributions of points that lie on a disc of radius 1.
- Suppose (R, Θ) are the polar coordinates of a random point.
- Suppose that $R \sim U(0, 1)$ and $\Theta \sim U(0, 2\pi)$.

What does this distribution look like? Is it uniform over the disc?

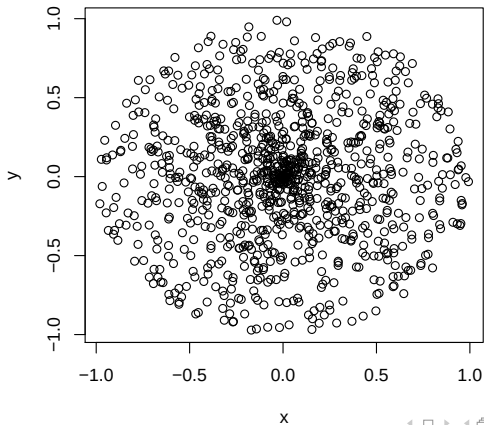
In other words, R and Θ are uniformly distributed... are X and Y ?

R code to sample distribution on disc

```
1 disc = function(n) {  
2   r = runif(n, 0, 1)  
3   theta = runif(n, 0, 2*pi)  
4   x = r*cos(theta)  
5   y = r*sin(theta)  
6   data.frame(x, y)  
7 }
```

Disc scatterplot

```
1 plot(disc(1000))
```



PDF for a bivariate transformation

Let X and Y have joint PDF $f_{X,Y}(x,y)$ Let $A \subseteq \mathbb{R}^2$ be the set

$$A = \{(x,y) : f_{X,Y}(x,y) > 0\}.$$

Suppose $T : A \rightarrow B$ is an invertible map where $B \subseteq \mathbb{R}^2$.

Then define random variables $(U, V) = T(X, Y)$.

Then the joint PDF of U, V is

$$f_{U,V}(u,v) = |J| f_{X,Y}(T^{-1}(u,v)) \quad \text{when } (u,v) \in B$$

where J is the determinant

$$J = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

If $(u,v) \notin B$ then $f_{U,V}(u,v) = 0$.

PDF for a bivariate transformation

Let X and Y have joint PDF $f_{X,Y}(x,y)$ Let $A \subseteq \mathbb{R}^2$ be the set

$$A = \{(x,y) : f_{X,Y}(x,y) > 0\}.$$

Suppose $T : A \rightarrow B$ is an invertible map where $B \subseteq \mathbb{R}^2$.

Then define random variables $(U, V) = T(X, Y)$.

Then the joint PDF of U, V is

$$f_{U,V}(u,v) = |J| f_{X,Y}(T^{-1}(u,v)) \quad \text{when } (u,v) \in B$$

where J is the determinant

$$J = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

If $(u,v) \notin B$ then $f_{U,V}(u,v) = 0$.

Transformation examples

Example

If (X, Y) are uniformly distributed on the unit disc, find the joint PDF of (R, Θ) where R and Θ are the corresponding polar coordinates. Also find the marginal PDFs for R and Θ .

X and Y are uniformly distributed... are R and Θ ?

Transformation examples

Example

If (X, Y) are uniformly distributed on the unit disc, find the joint PDF of (R, Θ) where R and Θ are the corresponding polar coordinates. Also find the marginal PDFs for R and Θ .

X and Y are uniformly distributed... are R and Θ ?

Solution

$$A = \text{disc with unit radius, } \{(x, y) : x^2 + y^2 \leq 1\}$$

$$B = \{(r, \theta) : 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}$$

T is the transformation

$$(x, y) \longrightarrow \left((x^2 + y^2)^{1/2}, \tan^{-1}(y/x) \right) = (r, \theta).$$

The inverse is

$$(r, \theta) \longrightarrow (r \cos \theta, r \sin \theta) = (x, y).$$

$$J = \det \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Solution

$$A = \text{disc with unit radius, } \{(x, y) : x^2 + y^2 \leq 1\}$$

$$B = \{(r, \theta) : 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}$$

T is the transformation

$$(x, y) \longrightarrow \left((x^2 + y^2)^{1/2}, \tan^{-1}(y/x) \right) = (r, \theta).$$

The inverse is

$$(r, \theta) \longrightarrow (r \cos \theta, r \sin \theta) = (x, y).$$

$$J = \det \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Solution

$$A = \text{disc with unit radius, } \{(x, y) : x^2 + y^2 \leq 1\}$$

$$B = \{(r, \theta) : 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}$$

T is the transformation

$$(x, y) \longrightarrow \left((x^2 + y^2)^{1/2}, \tan^{-1}(y/x) \right) = (r, \theta).$$

The inverse is

$$(r, \theta) \longrightarrow (r \cos \theta, r \sin \theta) = (x, y).$$

$$J = \det \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Solution

If X, Y are uniform on A then

$$f_{X,Y}(x,y) = \begin{cases} 1/\pi & (x,y) \in A \text{ disc has area } \pi; \\ 0 & \text{otherwise} \end{cases}$$

So

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(T^{-1}(r,\theta)) \times |J| \\ &= 1/\pi \times r. \end{aligned}$$

Therefore

$$\begin{aligned} f_{\Theta}(\theta) &= \int_{r=0}^{r=1} f_{R,\Theta}(r,\theta) dr \\ &= \left[\frac{r^2}{2\pi} \right]_{r=0}^{r=1} \\ &= \frac{1}{2\pi}, \end{aligned}$$

Thus $\Theta \sim U(0, 2\pi)$.

Solution

If X, Y are uniform on A then

$$f_{X,Y}(x,y) = \begin{cases} 1/\pi & (x,y) \in A \text{ disc has area } \pi; \\ 0 & \text{otherwise} \end{cases}$$

So

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(T^{-1}(r,\theta)) \times |J| \\ &= 1/\pi \times r. \end{aligned}$$

Therefore

$$\begin{aligned} f_{\Theta}(\theta) &= \int_{r=0}^{r=1} f_{R,\Theta}(r,\theta) dr \\ &= \left[\frac{r^2}{2\pi} \right]_{r=0}^{r=1} \\ &= \frac{1}{2\pi}, \end{aligned}$$

Thus $\Theta \sim U(0, 2\pi)$.

Solution

If X, Y are uniform on A then

$$f_{X,Y}(x,y) = \begin{cases} 1/\pi & (x,y) \in A \text{ disc has area } \pi; \\ 0 & \text{otherwise} \end{cases}$$

So

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(T^{-1}(r,\theta)) \times |J| \\ &= 1/\pi \times r. \end{aligned}$$

Therefore

$$\begin{aligned} f_{\Theta}(\theta) &= \int_{r=0}^{r=1} f_{R,\Theta}(r,\theta) dr \\ &= \left[\frac{r^2}{2\pi} \right]_{r=0}^{r=1} \\ &= \frac{1}{2\pi}, \end{aligned}$$

Thus $\Theta \sim U(0, 2\pi)$.

Solution

If X, Y are uniform on A then

$$f_{X,Y}(x,y) = \begin{cases} 1/\pi & (x,y) \in A \text{ disc has area } \pi; \\ 0 & \text{otherwise} \end{cases}$$

So

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(T^{-1}(r,\theta)) \times |J| \\ &= 1/\pi \times r. \end{aligned}$$

Therefore

$$\begin{aligned} f_{\Theta}(\theta) &= \int_{r=0}^{r=1} f_{R,\Theta}(r,\theta) dr \\ &= \left[\frac{r^2}{2\pi} \right]_{r=0}^{r=1} \\ &= \frac{1}{2\pi}, \end{aligned}$$

Thus $\Theta \sim U(0, 2\pi)$.

Solution

If X, Y are uniform on A then

$$f_{X,Y}(x,y) = \begin{cases} 1/\pi & (x,y) \in A \text{ disc has area } \pi; \\ 0 & \text{otherwise} \end{cases}$$

So

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(T^{-1}(r,\theta)) \times |J| \\ &= 1/\pi \times r. \end{aligned}$$

Therefore

$$\begin{aligned} f_{\Theta}(\theta) &= \int_{r=0}^{r=1} f_{R,\Theta}(r,\theta) dr \\ &= \left[\frac{r^2}{2\pi} \right]_{r=0}^{r=1} \\ &= \frac{1}{2\pi}, \end{aligned}$$

Thus $\Theta \sim U(0, 2\pi)$.

Solution

Now for the marginal for R :

$$\begin{aligned}f_R(r) &= \int_{\theta=0}^{\theta=2\pi} f_{R,\Theta}(r,\theta) d\theta \\&= \int_{\theta=0}^{\theta=2\pi} \frac{r}{\pi} d\theta \\&= \left[\frac{r}{\pi} \theta \right]_{\theta=0}^{\theta=2\pi} \\&= \frac{r}{\pi} 2\pi = 2r.\end{aligned}$$

Solution

Now for the marginal for R :

$$\begin{aligned}f_R(r) &= \int_{\theta=0}^{\theta=2\pi} f_{R,\Theta}(r, \theta) d\theta \\&= \int_{\theta=0}^{\theta=2\pi} \frac{r}{\pi} d\theta \\&= \left[\frac{r}{\pi} \theta \right]_{\theta=0}^{\theta=2\pi} \\&= \frac{r}{\pi} 2\pi = 2r.\end{aligned}$$

Solution

Now for the marginal for R :

$$\begin{aligned}f_R(r) &= \int_{\theta=0}^{\theta=2\pi} f_{R,\Theta}(r, \theta) d\theta \\&= \int_{\theta=0}^{\theta=2\pi} \frac{r}{\pi} d\theta \\&= \left[\frac{r}{\pi} \theta \right]_{\theta=0}^{\theta=2\pi} \\&= \frac{r}{\pi} 2\pi = 2r.\end{aligned}$$

Solution

Now for the marginal for R :

$$\begin{aligned}f_R(r) &= \int_{\theta=0}^{\theta=2\pi} f_{R,\Theta}(r, \theta) d\theta \\&= \int_{\theta=0}^{\theta=2\pi} \frac{r}{\pi} d\theta \\&= \left[\frac{r}{\pi} \theta \right]_{\theta=0}^{\theta=2\pi} \\&= \frac{r}{\pi} 2\pi = 2r.\end{aligned}$$

Transformation examples

Example

If X and Y are independent $\text{Exp}(1)$ random variables, and $U = X + Y$, $V = X/(X + Y)$, then find the joint distribution of U and V .

Solution

$$u = x + y, \quad v = \frac{x}{x + y} \quad \text{defines map } T.$$

So

$$A = (0, \infty) \times (0, \infty) \quad (x, y \text{ both positive})$$

$$B = (0, \infty) \times (0, 1) \quad (v \in (0, 1)).$$

$$T^{-1} : (u, v) \longrightarrow (uv, u(1 - v)) = (x, y).$$

$$J = \det \begin{pmatrix} v & u \\ 1 - v & -u \end{pmatrix} = -uv - u(1 - v) = -u.$$

Solution

$$u = x + y, \quad v = \frac{x}{x + y} \quad \text{defines map } T.$$

So

$$A = (0, \infty) \times (0, \infty) \quad (x, y \text{ both positive})$$

$$B = (0, \infty) \times (0, 1) \quad (V \in (0, 1)).$$

$$T^{-1} : (u, v) \longrightarrow (uv, u(1 - v)) = (x, y).$$

$$J = \det \begin{pmatrix} v & u \\ 1 - v & -u \end{pmatrix} = -uv - u(1 - v) = -u.$$

Solution

$$u = x + y, \quad v = \frac{x}{x + y} \quad \text{defines map } T.$$

So

$$A = (0, \infty) \times (0, \infty) \quad (x, y \text{ both positive})$$

$$B = (0, \infty) \times (0, 1) \quad (V \in (0, 1)).$$

$$T^{-1} : (u, v) \longrightarrow (uv, u(1 - v)) = (x, y).$$

$$J = \det \begin{pmatrix} v & u \\ 1 - v & -u \end{pmatrix} = -uv - u(1 - v) = -u.$$

Solution

$$u = x + y, \quad v = \frac{x}{x + y} \quad \text{defines map } T.$$

So

$$A = (0, \infty) \times (0, \infty) \quad (x, y \text{ both positive})$$

$$B = (0, \infty) \times (0, 1) \quad (V \in (0, 1)).$$

$$T^{-1} : (u, v) \longrightarrow (uv, u(1 - v)) = (x, y).$$

$$J = \det \begin{pmatrix} v & u \\ 1 - v & -u \end{pmatrix} = -uv - u(1 - v) = -u.$$

Solution

$$f_{X,Y}(x,y) = e^{-(x+y)} \quad (X, Y \sim \text{Exp}(1)).$$

So

$$\begin{aligned} f_{U,V}(u,v) &= |u| \times e^{-(uv+u-uv)} \\ &= ue^{-u} \quad \text{since } u \geq 0. \end{aligned}$$

Hence U and V are independent: joint PDF factorizes.

$$f_U(u) = ue^{-u} (0 \leq u < \infty) \quad \text{and} \quad V \sim U(0,1).$$

Solution

$$f_{X,Y}(x,y) = e^{-(x+y)} \quad (X, Y \sim \text{Exp}(1)).$$

So

$$\begin{aligned} f_{U,V}(u,v) &= |u| \times e^{-(uv+u-uv)} \\ &= ue^{-u} \quad \text{since } u \geq 0. \end{aligned}$$

Hence U and V are independent: joint PDF factorizes.

$$f_U(u) = ue^{-u} (0 \leq u < \infty) \quad \text{and} \quad V \sim U(0,1).$$

Solution

$$f_{X,Y}(x,y) = e^{-(x+y)} \quad (X, Y \sim \text{Exp}(1)).$$

So

$$\begin{aligned} f_{U,V}(u,v) &= |u| \times e^{-(uv+u-uv)} \\ &= ue^{-u} \quad \text{since } u \geq 0. \end{aligned}$$

Hence U and V are independent: joint PDF factorizes.

$$f_U(u) = ue^{-u} (0 \leq u < \infty) \quad \text{and} \quad V \sim U(0,1).$$

Solution

$$f_{X,Y}(x,y) = e^{-(x+y)} \quad (X, Y \sim \text{Exp}(1)).$$

So

$$\begin{aligned} f_{U,V}(u,v) &= |u| \times e^{-(uv+u-uv)} \\ &= ue^{-u} \quad \text{since } u \geq 0. \end{aligned}$$

Hence U and V are independent: joint PDF factorizes.

$$f_U(u) = ue^{-u} (0 \leq u < \infty) \quad \text{and} \quad V \sim U(0,1).$$

Lecture 5: The multivariate normal distribution

The bivariate normal distribution

Suppose $\mu_x, \mu_y, \sigma_x \geq 0, \sigma_y \geq 0$ 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 $\mathbb{R}^n \rightarrow \mathbb{R}$ 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

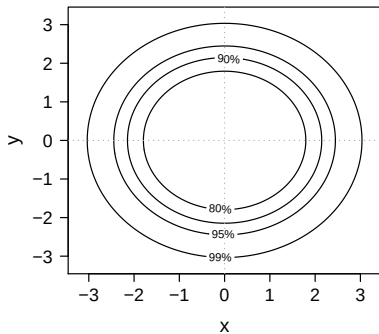
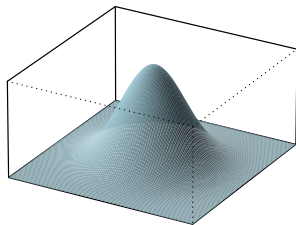
$$\underline{x}^T A \underline{x} \geq 0$$

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.

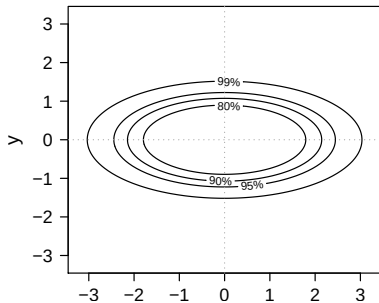
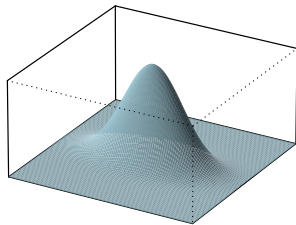
Pictures

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



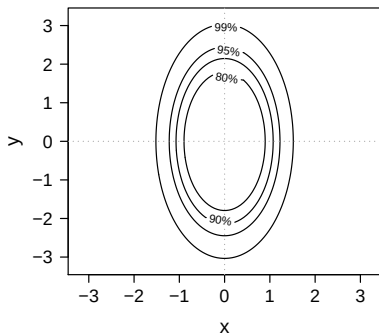
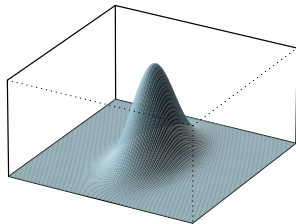
Pictures

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



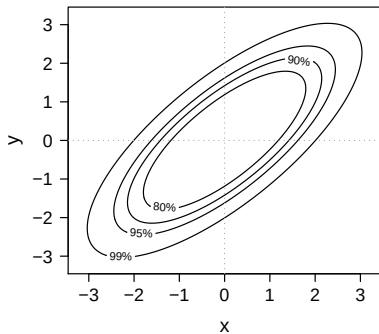
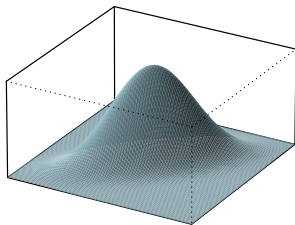
Pictures

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



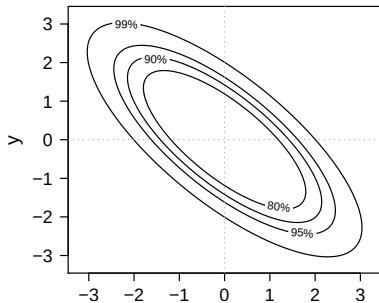
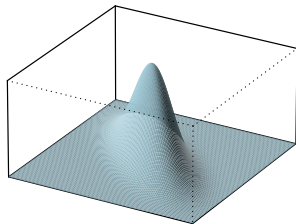
Pictures

$$\sigma_x = \sigma_y, \rho = 0.75$$



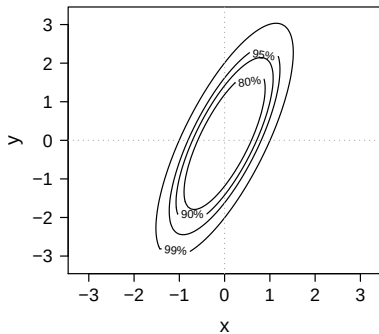
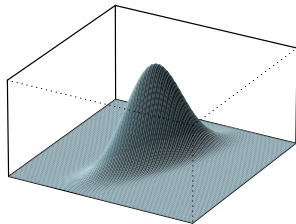
Pictures

$$\sigma_x = \sigma_y, \rho = -0.75$$



Pictures

$$2\sigma_x = \sigma_y, \rho = 0.75$$



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. $\text{Cor}(X, Y) = \rho$. 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 5.1

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 $\Pr(X^2 + Y^2 < 9)$.

Solution

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

Using the formula for the conditional

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

Example 5.1

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 $\Pr(X^2 + Y^2 < 9)$.

Solution

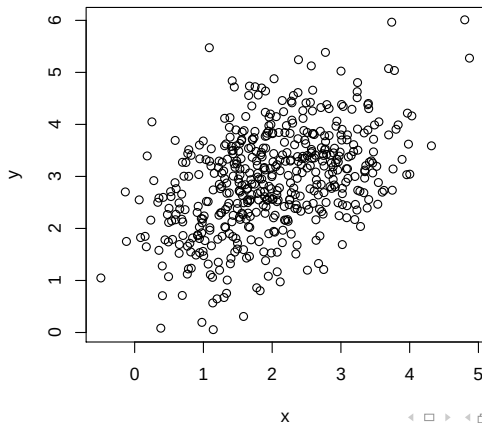
The marginal distribution of X is $X \sim N(2, 1^2)$.

Using the formula for the conditional

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

Simulation results

```
1 npts = 500
2 x = rnorm(npts, mean=2, sd = 1)
3 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:

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

Answer $\simeq 0.2776$

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 31, 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

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

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

Extra example

We have $Z = X + 3Y$. Using result 4 on page 31, 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

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

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

Extra example

We have $Z = X + 3Y$. Using result 4 on page 31, 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

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

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

Extra example

We have $Z = X + 3Y$. Using result 4 on page 31, we have

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

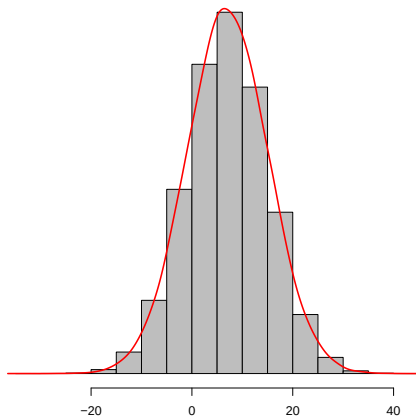
Now from the variance-covariance matrix, we have $\rho\sigma_x\sigma_y = 2$.

Thus

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

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

Extra example



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.

$$\underline{X} = (X_1, \dots, X_n)^T.$$

Then

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

if $X_1, \dots, X_n$ 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).