

Lecture: 2pm, Monday 21st October

MAS2906: Computational Probability and Statistics with R

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

Semester 1, 2024/25

Where we're at...

1. Three week R Refresher

- The basics: Basic operations, vectors and dataframes
- Graphical and numerical summaries
- Functions, for loops and if statements

2. Simulating random variables

- Using R's built-in functions for common discrete and continuous distributions
- More advanced probability models

Where we're at...

1. Three week R Refresher

- The basics: Basic operations, vectors and dataframes
- Graphical and numerical summaries
- Functions, for loops and if statements

2. Simulating random variables

- Using R's built-in functions for common discrete and continuous distributions
- More advanced probability models

Where we're at...

1. Three week R Refresher

- The basics: Basic operations, vectors and dataframes
- Graphical and numerical summaries
- Functions, for loops and if statements

2. Simulating random variables

- Using R's built-in functions for common discrete and continuous distributions
- **More advanced probability models**

Where we're at...

1. Three week R Refresher

- The basics: Basic operations, vectors and dataframes
- Graphical and numerical summaries
- Functions, for loops and if statements

2. Simulating random variables

- Using R's built-in functions for common discrete and continuous distributions
- More advanced probability models

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.9117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288389
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665289
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 3)$.	<code>pexp(3,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.9117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288389
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665289
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.9117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288389
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665289
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.9117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.9288389
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.9665289
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.9117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288389
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665289
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.9288385
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.9665285
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.9288305
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665255
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665255
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665255
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 3)$.	<code>dpois(3,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(3,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(2,0.4)</code>	0.16
	Find $\Pr(Z \leq 3)$.	<code>pgeom(3,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(3,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pggeom(3,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pggeom(3,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pggeom(3,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pggeom(3,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(3,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(3,0.4)</code>	0.784

Recap quiz

Model	Question	R code	Answer
$X \sim \text{Exp}(2)$	Find $\Pr(X < 2)$.	<code>pexp(2,2)</code>	0.9817
	Find $\Pr(X \leq 2)$.	<code>pexp(2,2)</code>	0.9817
	Generate 3 observations.	<code>rexp(3,2)</code>	0.8117, 1.0386, 0.0174
$Y \sim \text{Po}(4)$	Find $\Pr(Y < 5)$.	<code>ppois(4,4)</code>	0.6288369
	Find $\Pr(Y \geq 4)$.	<code>1-ppois(3,4)</code>	0.5665299
	Find $\Pr(Y = 2)$.	<code>dpois(2,4)</code>	0.1465251
$Z \sim \text{Geom}(0.4)$	Find $\Pr(Z = 2)$.	<code>dgeom(1,0.4)</code>	0.24
	Find $\Pr(Z \leq 3)$.	<code>pgeom(2,0.4)</code>	0.784
	Find $\Pr(Z < 4)$.	<code>pgeom(2,0.4)</code>	0.784

3.3 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 $\Pr(X > 30)$?

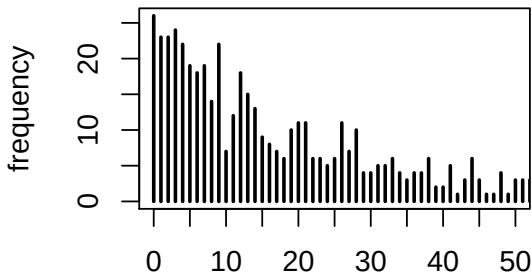
3.3 Arrival model – R code

```
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 }
```

3.3 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))
```



3.3 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.

3.3 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
```

3.4 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.

3.4 Mixtures of normal distributions

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

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

3.4 Mixtures of normal distributions

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

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

3.4 Mixtures of normal distributions

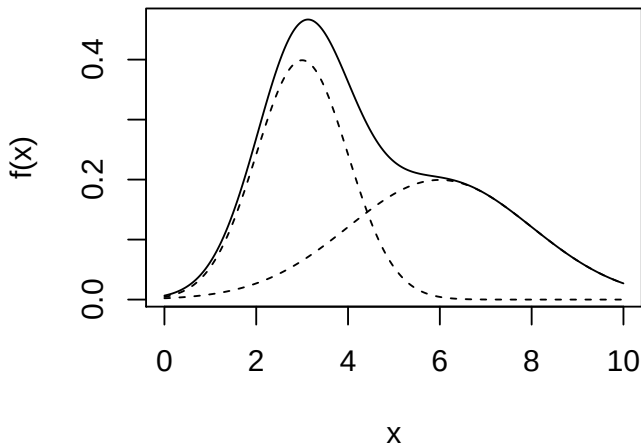
If a random variable X has PDF corresponding to $f(x)$ we say it is a **mixture** of normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ with weights w_1, w_2 .

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

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

3.4 Example

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



3.4 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 $N(\mu_2, \sigma_2^2)$.

3.4 Sampling from a mixture distribution

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

Example 3.4

Example

Investigations into infants suffering from *severe idiopathic respiratory distress syndrome* have shown that whether the infant survives may be related to their weight at birth.

Suppose that you are interested in developing a distribution for the birth weight X of such infants.

You have a normal $N(2.3, 0.52^2)$ distribution for the birth weight (in kg) of infants who survive and a normal $N(1.7, 0.66^2)$ distribution for infants who die.

Example 3.4

Example

- (a) If you believe that the proportion of infants that survive is 0.6, write down a suitable Normal mixture distribution for X .
- (b) Produce some R code to sample from the normal mixture distribution you propose in part (a).
- (c) Use your code in part (b) to simulate a sample of size 1000 from your normal mixture distribution, and produce a histogram of your sample.

Example 3.4: Solution

Using our notation from earlier, define

$$J = \begin{cases} 1 & \text{If the infant survives} \\ 2 & \text{If the infant dies} \end{cases}$$

Then

$$\Pr(J = 1) = 0.6 \quad (= w_1) \quad \text{and}$$

$$\Pr(J = 2) = 0.4 \quad (= w_2).$$

Example 3.4: Solution

Using our notation from earlier, define

$$J = \begin{cases} 1 & \text{If the infant survives} \\ 2 & \text{If the infant dies} \end{cases}$$

Then

$$\begin{aligned} \Pr(J = 1) &= 0.6 \quad (= w_1) \quad \text{and} \\ \Pr(J = 2) &= 0.4 \quad (= w_2). \end{aligned}$$

Example 3.4: Solution

We also know that:

- If $J = 1$, we need to sample from $N(2.3, 0.52^2)$
- If $J = 2$, we need to sample from $N(1.7, 0.66^2)$

So we have

$$0.6 \cdot N(2.3, 0.52^2) + 0.4 \cdot N(1.7, 0.66^2).$$

Example 3.4: Solution

We also know that:

- If $J = 1$, we need to sample from $N(2.3, 0.52^2)$
- If $J = 2$, we need to sample from $N(1.7, 0.66^2)$

So we have

$$0.6 \cdot N(2.3, 0.52^2) + 0.4 \cdot N(1.7, 0.66^2).$$

Example 3.4: Solution

Therefore we can use our `normal.mixture` function from earlier, with $\mu_1 = 2.3$, $\sigma_1 = 0.52$, $\mu_2 = 1.7$ and $\sigma_2 = 0.66$, giving:

```
test = normal.mixture(1000,2.3,0.52,0.6, 1.7,0.66,0.4)
hist(test)
```

Example 3.4: Solution

Therefore we can use our `normal.mixture` function from earlier, with $\mu_1 = 2.3$, $\sigma_1 = 0.52$, $\mu_2 = 1.7$ and $\sigma_2 = 0.66$, giving:

```
1 test = normal.mixture(1000,2.3,0.52,0.6, 1.7,0.66,0.4)
2 hist(test)
```

Example 3.4: Solution

