## WinBUGS : part 1

Bruno Boulanger Jonathan Jaeger Astrid Jullion Philippe Lambert





## Agenda

- Introduction to WinBUGS
- Exercice 1 : Normal with unknown mean and variance
  - Example description
  - Model specification in WinBUGS
  - Data and initial values in WinBUGS
  - Process in WinBUGS
  - Results in WinBUGS
- > Exercice 2 : Comparison of 2 means-unequal variance



WINBUGS
---------

#### > Free software :

http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

- WinBUGS implements various MCMC (Markov Chain Monte Carlo) algorithms to generate samples from posterior distributions.
- WinBUGS: Windows version of the BUGS program (Bayesian analysis using Gibbs Sampling).



## WinBUGS

Posterior of delta Odds ratio 20 S <del>6</del> 4 Density Density ജ က 2 2 6 0.06 0.04 0.08 0.10 1.3 1.5 1.7 1.9 delta = theta1 - theta2 gamma

Mean and quantiles of δ:
 [1] 0.066
 2.5% 5% 50% 95% 97.5%
 0.051 0.053 0.066 0.079 0.082

```
Mean and quantiles of γ:
[1] 1.574
2.5% 5% 50% 95% 97.5%
1.422 1.447 1.572 1.712 1.737
```

Identified posterior distributions are not needed in WinBUGS.

## **Example description**

- Solution Assume that it is known from large studies that the mean cholesterol level in children aged 2-14 is 175 mg/dL. It is desirable to determine if there is familial aggregation of cholesterol levels.
- A group of fathers who had heart attack and elevated cholesterol levels (250 mg/dL) was identified and the cholesterol level of their offspring aged 2-14 was measured.
- The cholesterol levels of such 100 children have mean
   207mg/dL and standard deviation 30 mg/dL.
- Solution Solution



## Example description

- By experience, it is known that the cholesterol levels are approximately normally distributed.
- Solution Therefore, we shall assume that  $Y_i \sim N(\mu; \sigma^2)$  (i = 1,...,100) where  $Y_i$  denotes the cholesterol level of the ith child.
- Precision  $\tau = 1/\sigma^2$



## Example description

• Prior for  $\mu$  :  $\mu \sim N(\mu_0, \sigma^2_0)$ 

A large variance  $\sigma^2_0$  would translate a large uncertainty on the values of  $\mu_0$ .

**Prior for**  $\tau$  :  $\tau \sim G(a,b)$  with mean a/b and variance a/b<sup>2</sup>

A large prior variance would translate uncertainty on the plausible values for  $\tau$ . A large prior variance can be obtained by taking a = b = 0.0001 (say).







> To specify a Bayesian model, we need:

• The likelihood:  $Y_i \sim N(\mu; \sigma^2)$  (i = 1,...,100)

We use:  $\tau = 1/\sigma^2$  (where  $\tau$  is the precision).

### • Prior distribution of the model parameters: $\mu$ ; $\tau$

• We consider non-informative prior distributions:  $\mu \sim N(200, 10^4)$  tau~ Gamma(0.0001,0.0001)



#### Some common univariate distributions in WinBUGS:

10

dbin	binomial	r ~ dbin(p,n)
dnorm	normal	x ~ dnorm(mu, <b>tau</b> )
dlnorm	log-normal	x ~ dlnorm(mu, <b>tau</b> )
dpois	poisson	r ~ dpois(lambda)
dunif	uniform	x ~ dunif(a,b)
dgamma	gamma	x ~ dgamma(a,b)
dbeta	beta	x ~ dbeta(a,b)



## Some rules :

- For loop: for (i in 1:n){ }
- Indice: y[i]
- Distribution: y[i] ~ dnorm(mu[i],tau)
- Assignment: variance <- 1/tau (computed in a deterministic way)</li>



## The likelihood:

for (i in 1:n){
 y[i]~dnorm(mu,tau)}

#### Prior distribution of the model parameters:

mu ~ dnorm (200 , 0.0001) tau ~ dgamma (0.0001, 0.0001)









## Data and initial values in WinBUGS





#### DATA

Make sure to provide all the data in the model specification.

```
Ifor (i in 1:n) €
```

y[i]~dnorm(mu,tau)

```
}
```

#Prior

```
mu~dnorm(200,0.0001)
tau~dgamma(0.0001,0.0001)
```

#Computation of the standard deviation sd<-1/sqrt(tau)



#### DATA

1.77079E+02, 2.49650E+02, 1.87267E+02, 2.02742E+02, 2.20882E+02, 1.98411E +02, 1.65022E+02, 2.30645E+02, 2.32637E+02, 2.28971E+02, 2.11241E+02, 1.54139E+02, 1.53285E+02, 1.96436E+02, 2.06299E+02, 1.93195E+02, 2.13863E +02, 1.47117E+02, 2.86469E+02, 1.91467E+02, 2.34219E+02, 1.71060E+02, 1.72452E+02, 1.77047E+02, 2.02564E+02, 2.53186E+02, 2.24139E+02, 1.95245E +02, 1.50038E+02, 2.21593E+02, 2.05437E+02, 1.76135E+02, 1.92493E+02, 2.09831E+02, 1.77484E+02, 1.99355E+02, 2.21080E+02, 2.15441E+02, 2.76884E +02, 2.10186E+02, 2.13041E+02, 2.17956E+02, 2.51468E+02, 2.25837E+02, 2.23581E+02, 2.59690E+02, 1.96986E+02, 2.30439E+02, 2.13505E+02, 1.70554E +02, 2.04125E+02, 2.21069E+02, 2.29274E+02, 2.55735E+02, 2.21125E+02, 2.28745E+02, 2.20769E+02, 2.12672E+02, 2.02715E+02, 2.01250E+02, 2.51791E +02, 2.17497E+02, 2.26841E+02, 1.95284E+02, 1.95802E+02, 1.89577E+02, 2.36446E+02, 1.79112E+02, 2.12935E+02, 1.93293E+02, 1.84941E+02, 1.48437E +02, 2.22016E+02, 2.22240E+02, 2.10760E+02, 1.93071E+02, 1.90897E+02, 1.69097E+02, 2.04470E+02, 1.68452E+02, 2.70150E+02, 2.18594E+02, 1.54905E +02, 2.16406E+02, 2.21260E+02, 2.06473E+02, 1.64567E+02, 2.61956E+02, 1.71259E+02, 2.11077E+02, 2.35920E+02, 2.18069E+02, 1.57796E+02, 2.22255E +02, 2.05355E+02), **n**=1.00000E+02)



#### Inital values

Make sure to provide initial values for all model parameters.

for (i in 1:n){
 y[i]~dnorm(mu,tau)
 }

**#Prior** 

```
mu~dnorm(200,0.0001)
tau~dgamma(0.0001,0.0001)
```

#Computation of the standard deviation
sd<-1/sqrt(tau)</pre>

#### ➢ Inital values

list(tau=1.00000E+00, mu=2.00000E+02)

18



## Process steps





## Process steps

- 1- check model
- 2- load data
- ➢ 3- compile model
- ♦ 4- load initial values
- 5- generate burn-in values
- 6- parameters to be monitored
- P- perform the sampling to generate samples from the posteriors
- 8- display results



## 1. Check model





#### > Highlight the word "model"

#### ➢ Go to Model -> Specification -> check model

🖥 File	Tools	Edit	Attributes	Info	Model	Inference	Options	СОМ
iode	11				Spec	cification		
					Upd	ate		
#Like	lihood	t L			Mon	itor Met		
tor	(i in 1 fil∼dn	:n){ iormi	(mu tau)		Save	e State		
y	լյա	}	(114,644)		See	d		
4Dria	r				Scrip	ot		
mu tau Con	' ∼dnor ∼dgar ∩putat ∝1/sq	rm(20 mma ion c rt(tai	00,0.000 i(0.0001, of the sta u)	1) 0.00 ndar	01) d devi	ation		
}								



😸 WinBUGS14 - [model\_mean]

nodel is syntactically correct

Stail

🕒 Inbox - ..

J 6

🚞 2 Wind... 👻

😂 2 Inter... 👻

3 Micr...

📓 File Tools Edit Attributes Info Model Inference Options COM Info Dev Tools Controls Doodle Map Obx Tut SQL Text Window Help



#### Specification Tool check model load data compile num of chains load inits for chain gen inits

# model is syntactically correct

Allocated Memory: 435'824 Bytes

👻 💽 Microsof...

WinBUG...

🦉 untitled ..

<u>....</u>

🎼 😫 🖸 🗿 🍰 🍢 🌮 💟 🔤 🤣 🏂 2:18 PM

🙎 2 Citri...

## 2. Load data





### Open your data file (*file -> open*) [.txt file]

#### Highlight the word "list" and click on "load data"

#### 🗱 WinBUGS14 - [data.txt]

🚡 File Tools Edit Attributes Info Model Inference Options COM Info Dev Tools Controls Doodle Map Obx Tut

**isti**v=c(2.51558E+02, 2.25099E+02, 1.66668E+02, 1.68022E+02, 2.19820E+02, 1.77079E+02. 2.49650E+02. 1.87267E+02. 2.02742E+02. 2.20882E+02. 1.98411E+02. 1.65022E+02, 2.30645E+02, 2.32637E+02, 2.28971E+02, 2.11241E+02, 1.54139E+02, 1.53285E+02. 1.96436E+02. 2.06299E+02. 1.93195E+02. 2.13863E+02. 1.47117E+02. 2.86469E+02, 1.91467E+02, 2.34219E+02, 1.71060E+02, 1.72452E+02, 1.77047E+02 2.02564E+02, 2.53186E+02, 2.24139E+02, 1.95245E+02, 1.50038E+02, 2.21593E+02 2.05437E+02. 1.76135E+02. 1.92493E+02. 2.09831E+02. 1.77484E+02. 1.99355E+02 2.21080E+02, 2.15441E+02, 2.76884E+02, 2.10186E+02, 2.13041E+02, 2.17956E+02, 2.51468E+02, 2.25837E+02, 2.23581E+02, 2.59690E+02, 1.96986E+02, 2.30439E+02 2.13505E+02. 1.70554E+02. 2.04125E+02. 2.21069E+02. 2.29274E+02. 2.55735E+02 2.21125E+02, 2.28745E+02, 2.20769E+02, 2.12672E+02, 2.02715E+02, 2.01250E+02 2.51791E+02.2.17497E+02.2.26841E+02.1.95284E+02.1.95802E+02.1.89577E+02. 2.36446E+02.1.79112E+02.2.12935E+02.1.93293E+02.1.84941E+02.1.48437E+02. 2.22016E+02, 2.22240E+02, 2.10760E+02, 1.93071E+02, 1.90897E+02, 1.69097E+02, 2.04470E+02. 1.68452E+02. 2.70150E+02. 2.18594E+02. 1.54905E+02. 2.16406E+02. 2.21260E+02, 2.06473E+02, 1.64567E+02, 2.61956E+02, 1.71259E+02, 2.11077E+02, 2.35920E+02, 2.18069E+02, 1.57796E+02, 2.22255E+02, 2.05355E+02), n=1.00000E+02)

😸 Specificat	tion Tool 🛛 🗙
check model	load data
compile	num of chains 1
load inits	for chain 1 📩
gen inits	





#### 🐺 WinBUGS14 - [data.txt]

🖹 File Tools Edit Attributes Info Model Inference Options COM Info Dev Tools Controls Doodle Map Obx Tut SQL Text Window Help

list(w=c(2.51558E+02, 2.25099E+02, 1.66668E+02, 1.68022E+02, 2.19820E+02, 1.77079E+02.2.49650E+02.1.87267E+02.2.02742E+02.2.20882E+02.1.98411E+02. 1.65022E+02. 2.30645E+02. 2.32637E+02. 2.28971E+02. 2.11241E+02. 1.54139E+02. 1.53285E+02, 1.96436E+02, 2.06299E+02, 1.93195E+02, 2.13863E+02, 1.47117E+02, 2.86469E+02. 1.91467E+02. 2.34219E+02. 1.71060E+02. 1.72452E+02. 1.77047E+02. 2.02564E+02. 2.53186E+02. 2.24139E+02. 1.95245E+02. 1.50038E+02. 2.21593E+02. 2.05437E+02, 1.76135E+02, 1.92493E+02, 2.09831E+02, 1.77484E+02, 1.99355E+02, 2.21080E+02, 2.15441E+02, 2.76884E+02, 2.10186E+02, 2.13041E+02, 2.17956E+02, 2.51468E+02. 2.25837E+02. 2.23581E+02. 2.59690E+02. 1.96986E+02. 2.30439E+02. 2.13505E+02. 1.70554E+02. 2.04125E+02. 2.21069E+02. 2.29274E+02. 2.55735E+02. 2.21125E+02, 2.28745E+02, 2.20769E+02, 2.12672E+02, 2.02715E+02, 2.01250E+02, 2.51791E+02, 2.17497E+02, 2.26841E+02, 1.95284E+02, 1.95802E+02, 1.89577E+02, 2.36446E+02.1.79112E+02.2.12935E+02.1.93293E+02.1.84941E+02.1.48437E+02. 2.22016E+02. 2.22240E+02. 2.10760E+02. 1.93071E+02. 1.90897E+02. 1.69097E+02. 2.04470E+02, 1.68452E+02, 2.70150E+02, 2.18594E+02, 1.54905E+02, 2.16406E+02, 2.21260E+02. 2.06473E+02. 1.64567E+02. 2.61956E+02. 1.71259E+02. 2.11077E+02. 2.35920E+02. 2.18069E+02. 1.57796E+02. 2.22255E+02. 2.05355E+02). n=1.00000E+02)



## 3. Compile





28

#### Solick on "compile"





#### 🐺 WinBUGS14 - [data.txt]

📓 File Tools Edit Attributes Info Model Inference Options COM Info Dev Tools Controls Doodle Map Obx Tut SQL Text Window Help

list(w=c(2.51558E+02, 2.25099E+02, 1.66668E+02, 1.68022E+02, 2.19820E+02, 1.77079E+02.2.49650E+02.1.87267E+02.2.02742E+02.2.20882E+02.1.98411E+02. 1.65022E+02. 2.30645E+02. 2.32637E+02. 2.28971E+02. 2.11241E+02. 1.54139E+02. 1.53285E+02. 1.96436E+02. 2.06299E+02. 1.93195E+02. 2.13863E+02. 1.47117E+02. 2.86469E+02, 1.91467E+02, 2.34219E+02, 1.71060E+02, 1.72452E+02, 1.77047E+02, 2.02564E+02. 2.53186E+02. 2.24139E+02. 1.95245E+02. 1.50038E+02. 2.21593E+02. 2.05437E+02. 1.76135E+02. 1.92493E+02. 2.09831E+02. 1.77484E+02. 1.99355E+02. 2.21080E+02, 2.15441E+02, 2.76884E+02, 2.10186E+02, 2.13041E+02, 2.17956E+02, 2.51468E+02, 2.25837E+02, 2.23581E+02, 2.59690E+02, 1.96986E+02, 2.30439E+02, 2.13505E+02. 1.70554E+02. 2.04125E+02. 2.21069E+02. 2.29274E+02. 2.55735E+02. 2.21125E+02, 2.28745E+02, 2.20769E+02, 2.12672E+02, 2.02715E+02, 2.01250E+02, 2.51791E+02, 2.17497E+02, 2.26841E+02, 1.95284E+02, 1.95802E+02, 1.89577E+02, 2.36446E+02.1.79112E+02.2.12935E+02.1.93293E+02.1.84941E+02.1.48437E+02. 2.22016E+02. 2.22240E+02. 2.10760E+02. 1.93071E+02. 1.90897E+02. 1.69097E+02. 2.04470E+02. 1.68452E+02. 2.70150E+02. 2.18594E+02. 1.54905E+02. 2.16406E+02. 2.21260E+02, 2.06473E+02, 1.64567E+02, 2.61956E+02, 1.71259E+02, 2.11077E+02, 2.35920E+02, 2.18069E+02, 1.57796E+02, 2.22255E+02, 2.05355E+02), n=1.00000E+02)



20

## 4. Load initial values











➢ If you want to run several chains, you have to specify initial values for each chain.

Solution Soluti Solution Solution Solution Solution Solution Solution S



# 5. Generate burn-in values





- ➢ Go into Model -> Update
- Specify the burn-in

#### Sclick on update

ucb

🐱 WinBUGS14	
ile Tools Edit Attributes Info Model Inference Options	COM Info Dev Tools Controls Doodle Map Ob
a model_mean	
model{	Specification Tool
# ikelihood	check model load data
for (i in 1:n){ y[i]~dnorm(mu,tau)	compile num of chains 2
, #Prior mu~dnorm(200,0.0001) tau~dgamma(0.0001,0.0001)	load inits for chain 2 🛃
#Computation of the standard deviation sd<-1/sqrt(tau)	updates 1000 refresh 100
}	update thin 1 iteration 0

35

# 6. Parameters to monitor





#### ➢ Go into Inference -> Samples



37



In "node", write the parameter you want to monitor and click on "set". **\_\_\_\_\_**38

- Repeat for all the parameters you want to monitor: mu, tau and sd.
- ➢ When done, write an asterisk

🗱 Sample Monitor T	iool 🛛 🚺
node ×	✓ chains 1 to 2 percentiles
beg 1 end	1000000 thin 1 2.5 5 10 25
	trace history density median 75
statscoda	quantiles bgr diag auto cor 95 97.5



7. Perform the sampling to generate samples from posteriors





ucb

In the "update tool", write the number of iterations and click on "update" 40

ile Tools Edit Attributes Info Model Inf	erence Options	COM Info	Dev To	ools Controls	Doodle	Мар	ОЬх	Tut S	5QL
🖥 model_mean									
model{	🙀 Upda	te Tool			रो				
#Likelihood	updates	5000	refresh	1000					
y[i]~dnorm(mu,tau)	update	thin 1	iteratio	n  6000					
}		over relax	🗌 adap	ting					
#Prior mu~dnorm(200,0.0001) tau~dgamma(0.0001,0.0001)									
#Computation of the standard dev sd<-1/sqrt(tau)	viation								
}									

## 8. display the results





ucb

In the "sample monitor tool", click on history to get the traces:

42

×	WinBUGS	514															
File	Tools E	dit At	tributes	Info	Model	Inference	Options	COM	Info	Dev	Tools	Controls	Doodle	Мар	Obx	Tut	SQL
			😹 Sar	nple	Monito	r Tool					×						
			node	×		•	<ul> <li>chains</li> </ul>	1 to	1	perce	entiles						
			beg [	1	enc	j 100000	0 thin	1		5 10 25							
			clear		set	trace	] histor	y d	ensity	mec 75	lian						
	Time seri	ies	stats		coda	quantiles	bgr dia	ag a	uto cor	90 95 <u>97.5</u>	5						
		п	nu												٦	^	200
	2	20.0	1								.					۳	JUU
	2	10.0	- http://	uhha	hall with	in the second second	a la pita fi	where he a	hilatu	and the	n, lifte	n hadada	իվորդիվել	11111			
	2	00.00	<b>H</b> Han	alit fa	an an talan	and ng di sa an	u kanalah j	- Julant	wph	li brit	and he	All & Halphin	n hard and	N. AN			
	1	90.0															
			1001		2000	D			4(	000				600	0		
								iteratio	on								
		10 0 S	d														
		35.0		اللب في	. Ոսվ	وبالمالية	يتفتأ المالي	միս, ետ	مىلىما		Lunan	International	باستار به ر				
		30.0		90 - 10 	and a state of the	na sa safisa.	alle fastante Alle fastante		diale at all of	ումահ	a matala	re on process an <b>la ba</b> da dia	n an an an	and the			
		25.0 -	and the second	יייויי	- 11 <b></b>	a di ka sa dalar da sa dalar da sa da s	alla, da Milda I	and a such	a contrata	A	ավեւտ	eller, e ok ete	an a	COLUMN T			
		20.0															
			1001		2000	D		iteratio	40 on	000				600	0		
		te	au														
	0	.002															
	0.0	0015	W.	L.	بالأر أعقران	والدراب للاطنان		لينزول الطبي	الاس	hul	فالعالد	STREES IN A DES	بالريادة الر	ur kille			
	0	.001	and the second	ير ليوا على		an an air air air an		in the state of the	i di statu	and add	, <b>Mah</b> lu	an hand had	والمرجة بالأرقاد	alka, Ju			
	5.0	0E-4		- "	11	er Heisen	n 91 - 11	1	1.1	1		a (1997) (f	11 T T	L1			
		_	1001		2000	D			40	000				600	0		
								iteratio	on								

▶ In the "sample monitor tool", click on "stats":

	nBU	GS14	4														
ile T	ools	Edit	Attrib	utes	Info	Model	Inference	Options	СОМ	Info	Dev	Tools	Controls	Doodle	Мар	ОЬх	Т
Sar	nple	. Mo	nitor T	ool					×								
ode	×				<b>•</b> 0	hains 1	to 1	percenti	les								
g [	1		end	1000	000	thin	1	2.5 5 10 25									
clear	r	SE	et	trac	e	history	density	25 median 75									
stats	:	CO	da 🕴	quant	iles	bgr diag	auto cor	90 95									
								97.5									
lo de	stat	tistio	cs														
lo de	stat de	tistic	cs	n .	od	M	C error'	2 5 %	medi	an 0	7 5 9/	ot	ort o	omple			^
lo de no	stat de	tistic	:s mea 207 1	n	<b>sd</b>	M	IC error:	<b>2.5%</b>	<b>medi</b> 207.1	an 9 2	<b>7.5%</b> 12 9	<b>st</b> a	art s	ample	,		^
lo de no mu sd	stai de	tistic	<b>mea</b> 207.1	n	<b>sd</b> 2.911 2.150	<b>M</b> 1 0.1	<b>IC error</b> 0432620 0311626	<b>2.5%</b> )1.5 } 1	<b>medi</b> 207.1 29.78	<b>an 9</b> 2	<b>7.5%</b> 12.9 4 47	<b>st</b> a 10	art s 01 5 01 5	ample	9		^

43



▶ In the "sample monitor tool", click on "density":

44



uch

Solution Solution



## EXERCICE 2: Comparison of 2 meansunequal variance





## Exercice 2

- The cholesterol level in 2 groups of children has been measured in a case-control study where the case group consists of  $n_1 =$ 100 children whose father suffered from cardiovascular diseases.
- The mean cholesterol levels of the  $n_2 = 74$  control kids is  $y_2 = 193.4$  with standard deviation  $s_2 = 17.3$ . In the case group, one has  $y_1 = 207.3$  and  $s_1 = 35.6$ . (Rosner 2000, p. 287).

The normality hypothesis is reasonable:

 $Y_{1i} \sim N(\mu_1, \sigma_1^2) \ (i = 1, ..., n_1)$ 

 $Y_{2i} \sim N(\mu_{2,}\sigma_{2}^{2}) (i = 1, ..., n_{2})$ 

Solution Estimate the difference between  $\mu_1$  and  $\mu_2$ :  $\delta = \mu_1 - \mu_2$ 

What do you conclude?



## Exercice 2

Non-informative priors :

- $\mu_1 \sim N(180, 10000)$
- $\mu_2 \sim N(180, 10000)$
- $\tau_1 \sim \text{Gamma}(0.0001, 0.0001)$
- $\tau_2 \sim \text{Gamma}(0.0001, 0.0001)$

```
with \tau_1 = 1/\sigma_1{}^2~ and \tau_2 = 1/\sigma_2{}^2~
```

Data : data.txt

Initial values : inits1.txt inits2.txt

