Typos from the 'Og

Christian P. Robert

October 10, 2011

Introducing Monte Carlo Methods with R

1. *(Thanks to Kazue Ishida, Japanese translator of the book)* The demos for chapters 2 and 5 do not work, due to an upgrade of R that invalidated my (much) older syntax. The demos should be fixed within the package mcsm any time soon.

2. *(Thanks to Jerry Sin)* On page 11, matrix summation in the matrix commands of Figure 1.2 should be matrix multiplication.

3. *(Thanks to Liaosa Xu from Virginia Tech)* On page 20, when we mention the uniform over the set

$$\left\{(a,b)\ :\ y_i(a+bx_i) > \log\frac{u_i}{1-u_i}\right\}$$

this set is missing (a) an intersection sign before the curly bracket and (b) a $(-1)^y_i$ instead of the y_i. It should be

$$\bigcap_{i=1}^{n}\left\{(a,b)\ :\ (-1)^{y_i}(a+bx_i) > \log\frac{u_i}{1-u_i}\right\}$$

4. *(Thanks to Matthieu Gomez)* In formula (2.1), the transform for the gamma $\mathcal{G})\alpha,\beta)$ distribution assumes β is a scale parameter, while the remainder of the book takes the opposite convention.

5. In Exercise 2.17, page 58, question d. should be *d. Show that the maximum of $b^{-a}(1-b)^{a-\alpha}$ is attained at $b = a/\alpha$.*

6. In Exercise 2.21, page 59, in item (ii), $||\theta||$ should be replaced by $\sqrt{\lambda}$ and question b. should be removed.

7. On page 71, due to the late inclusion of an extra-exercise in the book, the *above exercise* in Exercise 3.5 actually means Exercise 3.3.

8. *(Thanks to Brad McNeney, Simon Fraser University)* The end of Example 3.6 (page 75) is missing a marginal estimate, i.e. there is a $x(1-x)$ missing from $m(x)$. It should have been obvious from the estimates we derived, 19 and 16, which do not even appear on the support of the posterior distribution represented on Figure 3.5. The R code is given as

```
> mean(y[,1]*apply(y,1,f)/den)/mean(apply(y,1,h)/den)
[1] 19.33745
> mean(y[,2]*apply(y,1,f)/den)/mean(apply(y,1,h)/den)
[1] 16.54468
```

and should have been view source

```
> mean(y[,1]*f(y)/den)/mean(f(y)/den)
[1] 94.08314
> mean(y[,2]*f(y)/den)/mean(f(y)/den)
[1] 80.42832
```

A similar modification applies to the remark after eqn. (3.7) (page 76):

```
mean(apply(y,1,h)/den)
```

should be

```
mean(f(y)/den)
```

9. In Exercise 3.11, page 86, question c, a line got commented by mistake in the LaTeXfile and it should read *Explore the gain in efficiency from this method. Take* $a = 4.5$ *in part (a) and run an experiment to determine how many normal* $\mathcal{N}(0,1)$ *random variables would be needed to calculate* $P(Z > 4.5)$ *to the same accuracy obtained from using* 100 *random variables in this importance sampler.*

10. *(Thanks to Edward Kao, University of Houston)* In Exercise 3.17, page 88, question b, it should be $X|Y \sim \mathcal{G}a(1, y)$, not $X|Y \sim \mathcal{G}a(y, 1)$.

11. *(Thanks to Kazue Ishida, Japanese translator of the book)* In Example 4.4, the R code cannot run as provided and contains several mistakes.

    (a) in the remark of page 98, the second line of code should be

    ```
    > wachd[wachd<10^(-10)]=10^(-10)
    ```

    in order to eliminate all zeroes

    (b) in the code page 100, the negative log-perplexities should be put to zero:

    ```
    > plex[plex>0]=0
    > plech[plech>0]=0
    ```

    (this should come right after their definition)

    (c) till on page 100, here is a negative sign missing in the definition of the perplexities, `plob` and `ploch`: it should read

    ```
    > plob=apply(exp(-plex),1,quantile,c(.025,.975))
    > ploch=apply(exp(-plech),1,quantile,c(.025,.975))
    ```

    (d) Figure 4.4 (right) is wrong and needs to be redrawn by

```
> plot(1:10^4,0*(1:10^4),col="white",ylim=c(min(rbind(echbo,essbo)),max(rbind(e
> polygon(c(1:10^4,(10^4):1),c(essbo[1,],rev(essbo[2,])),col="grey65",border=FA
> polygon(c(1:10^4,(10^4):1),c(echbo[1,],rev(echbo[2,])),col="wheat",border=FALS
> plot(1:10^4,0*(1:10^4),col="white",ylim=c(min(rbind(plob,ploch)),max(rbind(pl
> polygon(c(1:10^4,(10^4):1),c(plob[1,],rev(plob[2,])),col="grey65",border=FALSE
> polygon(c(1:10^4,(10^4):1),c(ploch[1,],rev(ploch[2,])),col="wheat",border=FALS
```

12. *(Thanks to Ashley)* On page 128, running

```
> for (i in 1:(nlm(like,sta)$it)){
+ mmu=rbind(mmu,nlm(like,sta,iter=i)$est)}
```

produces the error message

```
Error in f(x, ?) : unused argument(s) (iter = 1)
```

as the latest versions of R, e.g. 2.12.1, calling nlm requires using `iterlim` instead of the abbreviation `iterr`:

```
> args(nlm)
function (f, p, ..., hessian = FALSE, typsize = rep(1, length(p)),
 fscale = 1, print.level = 0, ndigit = 12, gradtol = 1e-06,
 stepmax = max(1000 * sqrt(sum((p/typsize)^2)), 1000), steptol = 1e-06,
 iterlim = 100, check.analyticals = TRUE)
```

In addition, the function `nlm` produces the minimum of the first argument f and `like` should thus be defined as

```
> like=function(mu){
+   -sum(log((.25*dnorm(da-mu[1])+.75*dnorm(da-mu[2]))))}
```

to end up with local maxima as on Figure 5.2.

13. *(Thanks to Ashley)* Another mistake on page 128, namely that we used

```
> da=rbind(rnorm(10^2),2.5+rnorm(3*10^2))
```

instead of

```
> da=c(rnorm(10^2),2.5+rnorm(3*10^2))
```

to create a sample. Since both two normal samples have different sizes, `rbind` induces a duplication of the smaller sample, not what we intended!

14. In Exercise 4.5, page 101, the $\overline{X}_k$ should not be in bold fonts.

15. In Exercise 4.9, page 108, I commented too many lines when revising and thus the variance terms vanished. It should read

$$\begin{aligned}\mathbb{E}\left[\exp -X^2|y\right] &= \frac{1}{\sqrt{2\pi\sigma^2/y}}\int \exp\{-x^2\}\,\exp\{-(x-\mu)^2 y/2\sigma^2\}\,\mathrm{d}x\\ &= \frac{1}{\sqrt{2\sigma^2/y+1}}\exp\left\{-\frac{\mu^2}{1+2\sigma^2/y}\right\}\end{aligned}$$

16. In Exercise 4.13, page 122, following the removal of one exercise, Exercise 4.2 should read Exercise 4.1

17. In Exercise 4.15, page 122, $Bin(y)$ should be $Bin(n, y)$ (as in Problem 4.5 of Monte Carlo Statistical Methods)

18. *(Thanks to Ashley)* In Example 5.2, page 128, to simulate a sample of 400 observations from the mixture, we use

```
> da=rbind(rnorm(10^2),2.5+rnorm(3*10^2))
```

and need to replace `rbind` with +c+:

```
> da=sample(c(rnorm(10^2),2.5+rnorm(3*10^2)))
```

Note that `da` is not a sample from a mixture, because the proportion from each component is fixed.

19. On page 131, Exercise 5.3 has no simple encompassing set and the constraint should be replaced by

$$x^2(1+\sin(y/3)\cos(8x)) + y^2(2+\cos(5x)\cos(8y)) \le 1$$

20. *(Thanks to Ashley)* In Example 5.5, page 134, in the R code

```
+ apply(z,1,mean)}
```

should be

```
+ apply(z,1,prod)}
```

as we are not working with log-densities.

21. (cont'ed) Example 5.5 (pages 133–135) is completely off-the-mark! Not only the representation of the likelihood should have used `prod` instead of `mean`, not only the constant should call the `val` argument of `integrate`, not only `integrate` uses `lower` and `upper` rather than `from` and `to`, but also the approximation is missing a scale factor of 10, squared root of the sample size. The corrected R code is thus

```
> cau=rcauchy(10^2)
> mcau=median(cau)
> rcau=diff(quantile(cau,c(.25,.75)))/sqrt(10^2)
> f=function(x){
+   z=dcauchy(outer(x,cau,FUN="-"))
+   apply(z,1,prod)}
> fcst=integrate(f,lower=-20,upper=20)$val
> ft=function(x){f(x)/fcst}
> g=function(x){dt((x-mcau)/rcau,df=49)/rcau}
> curve(ft,from=-1,to=1,xlab="",ylab="",lwd=2)
> curve(g,add=T,lty=2,col="steelblue",lwd=2)
```

and Figure 5.5 (page 134) is therefore modified. Note that the fit by the tt distribution is not as perfect as before. A normal approximation would do better.

22. *(Thanks to Gilles Guillot, Technical University of Denmark)* On page 137, second equation from bottom

$$h(\theta+\beta\zeta)-h(\theta+\beta\zeta)$$

should be

$$h(\theta+\beta\zeta)-h(\theta-\beta\zeta)$$

23. *(Thanks to Gilles Guillot, Technical University of Denmark)* On page 138, Example 5.7, the denominator in the gradient should be `2*beta` [the error actually occurs twice. And once again in the R code.] In addition, the first paragraph of page 138 is missing details: the conditions on α and β are necessary and sufficient. Furthermore, `demo(Chapter.5)` triggers an error message [true, the shortcut max=TRUE instead of maximise=TRUE in optimise does not work with the latest versions of R] This is what occurs with version 2.11.1:

```
demo(Chapter.5)
---------------

Type  <Return>   to start :

> # Section 5.1, Numerical approximation
>
> ref=rcauchy(5001)

> f=function(y){-sum(log(1+(x-y)^2))}

> mi=NULL

> for (i in 1:400){
```

```
+   x=ref[1:i]
+   aut=optimise(f,interval=c(-10,10),max=T)
+   mi=c(mi,aut$max)
+   }
Error in f(arg, ?) : unused argument(s) (max = TRUE)

> optimise(f,interval=c(-10,10),maximum=T)
$maximum
[1] -2.571893

$objective
[1] -6.661338e-15
```

24. *(Thanks to Robin Ryder, CREST)* In Example 5.14, page 153, the sentence between parentheses should end up with *equal to 0*. And *step* should be replaced with *iteration* in the first paragraph.

25. *(Thanks to Robin Ryder, CREST)* On page 156, in Example 5.15, *likelihood surface* should be *log-likelihood surface*;

26. On page 158, Example 5.16 has a typo in that the EM sequence should be

$$\hat{\theta}_1 = \left\{ \frac{\theta_0\, x_1}{2+\theta_0} + x_4 \right\} \Big/ \left\{ \frac{\theta_0\, x_1}{2+\theta_0} + x_2 + x_3 + x_4 \right\}$$

    And *or, equivalently, by* should be *or, equivalently, with*;

27. On page 162, in the caption of Figure 5.14, *LE estimator* should be *MLE*;

28. In Exercise 5.15, page 163, the Z's in the formula should be in capital letters, namely

$$P(Z_i = 1) = 1 - P(Z_i = 2)$$

29. *(Thanks to Pierre Jacob and Robin Ryder, CREST)* In Exercise 5.17, page 164, the function ϕ should be written φ to be coherent with the notation of Example 5.14

30. Exercise 5.21, page 165, should have been removed as it duplicates Exercise 5.11 (but this redundancy is not going to confuse anyone!)

31. *(Thanks to Boris Vasiliev, Ottawa)* In Exercise 6.4, page 176, the second parameter of the gamma distribution $\mathcal{G}(a, b)$ is a rate parameter and not a scale parameter.

32. *(Thanks to Pierre Jacob and Robin Ryder, CREST)* In Exercise 6.13, page 197, both α and β must use a double exponential proposal in the Metropolis-Hastings algorithm of question b.

33. *(Thanks to Pierre Jacob and Robin Ryder, CREST)* In Exercise 6.15, page 197, the $\mathcal{L}(0, \omega)$ distribution should be a $\mathcal{N}(0, \omega)$ normal distribution,

34. In Example 7.3, page 204, part of the code is wrong: it should be

```
> sigma2=theta=rep(0,Nsim)                    #init arrays
> sigma2[1]=1/rgamma(1,shape=a,rate=b)        #init chains
> B=sigma2[1]/(sigma2[1]+n*tau2)
> theta[1]=rnorm(1,m=B*theta0+(1-B)*xbar,sd=sqrt(tau2*B))
```

instead of

```
> sigma=theta=rep(0,Nsim)                    #init arrays
> sigma{1}=1/rgamma(1,shape=a,rate=b)        #init chains
> B=sigma2{1}/(sigma2{1}+n*tau2)
> theta{1}=rnorm(1,m=B*theta0+(1-B)*xbar,sd=sqrt(tau2*B))
```

35. *(Thanks to Motohiro Ishida, Japanese translator of the book)* Always about Example 7.3, there is a missing line after the definition of the data, namely

```
> x=log(x)
```

that induces removing corrections I had made to compensate for inconsistencies (!):

(a) there should be no mention of $\log(\theta)$ nor of $\log(\sigma)$ in this example, either in Figure 7.2 or at the end of Example 7.3. The histograms are clearly for the mean and the standard deviation, not for the "log-mean and logstandard deviation".

(b) the posterior means of θ and σ^2 are 6.62 and 0.661, respectively, leading to an estimate of σ of 0.802.

36. In Exercise 7.3, there should be not log in the marginals of θ and σ^2.

37. In Algorithm 8, page 206, there are two commas before *given*;

38. in Example 7.6, page 210, I forgot to include the truncation probability $\Phi(a-\theta)^{n-m}$ in the likelihood and the notations are not completely coherent with Example 5.13 and 5.14 in that the x's became y's.

39. In the caption of Figure 7.6, page 2114, these are the allele probabilities, not the genotype probabilities;

40. in Exercise 7.21, page 235, `rtnorm` is missing `sigma` as one of its arguments.

41. In Exercise 7.22 c, page 235, the matrix is positive definite if and only if the condition is satisfied;

42. Exercise 7.23, page 235, has nothing wrong per se but it is rather a formal (mathematical) exercise.

43. in Exercise 7.25, page 235, the $x + a$ in question a should be $(x + a)$ to avoid any confusion.

44. On page 239, *the slower chain* should be *the slowest chain*;

45. On page 241, *you can load all* `coda` *functions* instead of *download*;

46. In the caption of Figure 8.2, page 244, the upper quantile is a 97.5

47. *(Thanks to Thomas Clerc, Fribourg)* On page 247, the complex number ι should be defined as the squared root of -1 instead of 1!

48. In Exercise 8.1, page 241, we compare an estimator δ_1 with δ_k, not δ_2. We need a bit more stability in the Markov chain to ensure that it has a finite variance. The assumption that both first moments are finite is not enough. Thomas Clerc from Fribourg also pointed out that the lazy R programming of the form `beta=c(beta,betaprop)` should not be encouraged!

49. In Example 8.9, page 264, $Y_k \sim \mathcal{N}(\theta_i, \sigma^2)$ should be $Y_i \sim \mathcal{N}(\theta_i, \sigma^2)$ and the θ_i's are normal, not the μ_i's.

50. In Exercise 8.7, page 265, it is the distribution on α that is closed-form, not the one on μ.

51. In Exercise 8.11, page 268, the y_i at the end of page 267 should be d_i;

52. In Exercise 8.16, page 268, jpeg is mistakenly qualified as open.