Rousseau Mathieu, 67001800

UCLouvain
Belgium
??/??/????

1 Introduction

Cancer cell growth has been studied in vitro in 10 parallel experiments. We have a collection of data $\mathcal{D}_i$ where $i = 1, \ldots, 10$ is the ith-experiment and where each $\mathcal{D}_i$ is a time series ranging from $t = 10$ to $t = 500$.

We describe the evolution of the tumor over time by a time-dependent Poisson distribution,

$$y(t) \sim \text{Pois}(\mu(t)) \tag{1}$$

where,

$$\mu(t) = \beta_0 \exp\left(\frac{\beta_1}{\beta_2}(1 - e^{-\beta_2 t})\right) \tag{2}$$

with $\beta_k > 0 (k \in \{0, 1, 2\})$.

Notice we can reparametrize this model as,

$$\mu(t) = \alpha_0 \exp(-\alpha_1 e^{-\alpha_2 t}) \tag{3}$$

with $\alpha_k > 0 (k \in \{0, 1, 2\})$.

Questions

Question 1

a) Examine $\mu(t)$ and its relative change over time $\frac{1}{\mu(t)}\frac{\mathrm{d}\mu(t)}{\mathrm{d}t}$ in order to provide an interpretation for each parameter of (**??**).

Let's compute the derivative first,

$$\begin{aligned}
\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} &= \mu(t)\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\beta_1}{\beta_2}(1 - e^{-\beta_2 t})\right) \\
&= \frac{\beta_1}{\beta_2}\mu(t)e^{-\beta_2 t}\beta_2 \\
&= \beta_1\mu(t)e^{-\beta_2 t}
\end{aligned}$$

We then have,

$$\frac{1}{\mu(t)}\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} = \beta_1 e^{-\beta_2 t}$$

b) Provide the formulas connecting the β_k and α_k parameters.

Let's compute the relative change over time of **??**,

$$\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} = \mu(t)\alpha_1\alpha_2 e^{-\alpha_2 t}$$

We then have,

$$\frac{1}{\mu(t)}\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} = \alpha_1\alpha_2 e^{-\alpha_2 t}$$

Comparing with the two equations as well as the relative changes over time,

$$\alpha_1 \alpha_2 e^{-\alpha_2 t} = \beta_1 e^{-\beta_2 t}$$

$$\beta_0 \exp\left(\frac{\beta_1}{\beta_2}(1 - e^{-\beta_2 t})\right) = \alpha_0 \exp(-\alpha_1 e^{-\alpha_2 t})$$

Question 2

a) Assuming independence between the data $\mathcal{D}_i$ collected during the experiment i where $\mathcal{D}_i = \{ y_i(t_j) : j = 1, \ldots, J = 99 \}$ with $y_i(t_j) \sim \text{Pois}(\mu(t_j))$, provide an analytic form for the likelihood function $L(\vec{\alpha}|\mathcal{D}_i)$ for experiment i.

We know that the number of cancer cells $y_{i,j}$ for the experiment i at any given time t_j, $j \in \{ 10, 15, 20, ..., 500 \}$ follow a Poisson distribution of parameter μ. Its probability mass function is given by,

$$p(y_{i,j}) \propto \mu^{y_{i,j}} \cdot e^{-\mu}$$

where an expression for μ is given above (**??**) as a function of $\alpha_k (k = 0, 1, 2)$.

Assuming independance, we then have,

$$\begin{aligned} L(\vec{\alpha}|\mathcal{D}_i) &= L((\alpha_0, \alpha_1, \alpha_2)|y_i) \\ &\propto \prod_{j=1}^{99} p(y_{i,j})^{y_{i,j}} \\ &\propto \prod_{j=1}^{99} \left(\mu^{y_{i,j}} \cdot e^{-\mu}\right)^{y_{i,j}} \end{aligned}$$

b) Provide a R function enabling to compute the log-likelihood for given parameter values $\theta_k = \log(\alpha_k)$ and data D_i.

We have,

$$\begin{aligned} l(\vec{\alpha}|\mathcal{D}_i) &:= \ln(L((\alpha_0, \alpha_1, \alpha_2)|y_i)) \\ &\propto \sum_{j=1}^{99} \ln\left(\left[\mu^{y_{i,j}} \cdot e^{-\mu}\right]^{y_{i,j}}\right) \\ &= \sum_{j=1}^{99} y_{i,j} \cdot \ln(\mu^{y_{i,j}} \cdot e^{-\mu}) \end{aligned}$$

```
mu <- function(alpha, t) {
  - alpha[1] * exp(- alpha[2] * exp(- alpha[3] * t))
}

log_likelihood <- function(theta, t, experiment) {
```

```
  n <- length(experiment)
  alpha <- exp(theta)

  result <- sum(sapply(1:n, function(j) {
    experiment[j] * log(mu(alpha = alpha, t = t[j])^experiment[j] * exp(-mu(alpha = alpha, t = t[j]))
  }))

  return(result)
}
```

c) Provide a R function enabling to compute the log-posterior $p(\vec{\theta}|\mathcal{D}_i)$ for given $\theta = \theta_0, \theta_1, \theta_2$, data $\mathcal{D}_i$ and large variance priors.

Since we have no information on the proportion of any parameter θ_k, we assume an uniform distribution,

$$\theta_k \sim \text{Unif}(0,1) \tag{4}$$

We find then,

$$\begin{aligned} p(\vec{\theta}|\mathcal{D}_i) &\propto \ln(L(\vec{\theta}|\mathcal{D}_i) \cdot (P(\theta_0)P(\theta_1)P(\theta_2))) \\ &\propto l(\vec{\theta}|\mathcal{D}_i) + \ln(P(\theta_0)) + \ln(P(\theta_1)) + \ln(P(\theta_2)) \\ &\propto \sum_{j=1}^{99} y_{i,j} \cdot \ln(\mu^{y_{i,j}} \cdot e^{-\mu}) + \ln(1_{0,1}) + \ln(1_{0,1}) + \ln(1_{0,1}) \end{aligned}$$

```
prior <- function(theta) {
  dunif(theta, 0, 1, log = TRUE)
}

log_posterior <- function(theta, t, experiment) {
  log_likelihood(theta, t, experiment) + prior(theta[0]) + prior(theta[1]) + prior(theta[2])
}
```

d) Using the preceding R function, obtain the mean vector and variance-covariance matrix in the Laplace approximation to the marginal posterior $\vec{\theta}$ given the data from experiment 1.

```
laplace_approximation <- function(log_posterior, inits, n_samples, ...) {
  fit <- optim(
    par = inits,
    fn = log_posterior,
    control = list(fnscale = -1),
    hessian = TRUE,
    ...
  )

  mean <- fit$param
  var_cov_matrix <- solve(-fit$hessian)

  return(list(
    mean = mean,
    var_cov_matrix = var_cov_matrix
  ))
}
```

```r
inits <- c(theta0 = 0, theta1 = 0, theta2 = 0)
lapprox <- laplace_approximation(log_posterior, inits, 10000, t = df$day, experiment = df$Exp1)

lapprox$mean
lapprox$var_cov_matrix
```

Question 3

a) Using a **Metropolis algorithm** (*with componentwise proposals*) and the R software (without using JAGS or other specific packages), draw a random sample from the posterior distribution of $(\vec{\theta}|\mathcal{D}_1)$.

Appendix