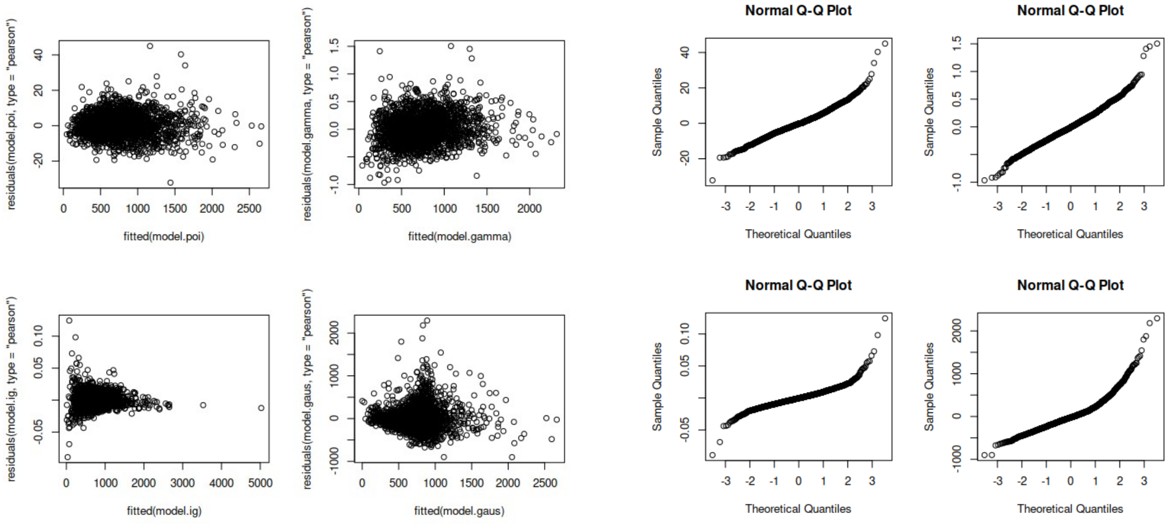
Statistical Modelling 2

Exercises 2

1.

a) Firstly, I selected to use log link since cd4 is always positive and other link functions caused some errors. After fitting the models, I obtained following residual plots



Based on the above plots, it seems that Gamma model is the preferred one since it’s residuals are quite well behaving and relatively small.

```
# b)
> new.data = data.frame(time=2.02, drugs=1, age=13.72)
> mu.hat = predict(model.gamma, newdata=new.data,
re.form=NA, type='response')
> mu.hat
       1
541.8597

# c)
> lb = quantile(yf.star, 0.1)
> ub = quantile(yf.star, 0.9)
> lb
    10%
522.424
> ub
     90%
1152.824
```

2.

```
# a)
> beta1.hat = summary(model.bin)$coef[2]
> beta1.hat
[1] 1.703157

# b)
> new.data = data.frame(id=24, sex=0, time=1.35, feed=0)
> mu.pred = predict(model.bin, newdata=new.data,
type='response')
> mu.pred
        1
0.7891318

# c)
> p.val = anova(model.H0, model.bin)$"Pr(>Chisq)"[2]
> p.val
[1] 4.81858e-09
X_2 is statistically significant

# d)
> cov.bi
          [,1]      [,2]
[1,]  1.298861 -1.127345
[2,] -1.127345  2.385338

# e)
> lb = YS.pred-z*sqrt(var.error)
> ub = YS.pred+z*sqrt(var.error)
> lb
       1
68.41443
> ub
       1
80.15473
```

3.

a)

Let $g(\mu_i) = \beta_0 + b_h$, where $b_h \sim N(0, \sigma_z^2)$

i) Calculate $E(y_i)$, when $\mu_i = \beta_0 + b_h$

$E(y_i) = E(\mu_i) = E(\beta_0 + b_h) = \beta_0 + E(b_h) = \beta_0$

ii) Calculate $E(y_i)$, when $\mu_i = exp(\beta_0 + b_h)$

$E(y_i) = E(\mu_i) = E(exp(\beta_0 + b_h)) = E(exp(\beta_0)exp(b_h)) = exp(\beta_0)E(exp(b_h))$

$E(exp(b_h))$ can be considered as the moment generating function of b_h. The moment function of normally distributed rv. x is defined as

$M_X(t) = E(e^{tX}) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$

Therefore

$E(e^{b_h}) = M_{b_h}(1) = exp(\frac{1}{2}\sigma_z^2)$

and

$E(y_i) = exp(\beta_0)E(exp(b_h)) = exp(\beta_0)exp(\frac{1}{2}\sigma_z^2)$

b)

Let $g(\mu_i) = \mathbf{x_i}'\boldsymbol{\beta} + b_h$, where $\mathbf{b} = \{b_h\} = (b_1, ..., b_q)' \sim N(\mathbf{0}, \sigma_z^2\mathbf{I})$

Write log likelihood of $\mathbf{b}$ as simply as possible.

$$
\begin{aligned}
f(\mathbf{b}) &= \frac{1}{\sqrt{(2\pi)^q det(\sigma_z^2\mathbf{I})}} exp(-\frac{1}{2}\mathbf{b}'(\sigma_z^2\mathbf{I})^{-1}\mathbf{b}) \\
&= (2\pi)^{-\frac{q}{2}}(\sigma_z^2)^{-\frac{q}{2}} exp(-\frac{1}{2}\mathbf{b}'(\sigma_z^2\mathbf{I})^{-1}\mathbf{b}) \\
&= (2\pi)^{-\frac{q}{2}}(\sigma_z^2)^{-\frac{q}{2}} exp(-\frac{1}{2\sigma_z^2}\mathbf{b}'\mathbf{b})
\end{aligned}
$$

$$
\begin{aligned}
l(\mathbf{b}) &= log(f(\mathbf{b})) \\
&= log((2\pi\sigma_z^2)^{-\frac{q}{2}} exp(-\frac{1}{2\sigma_z^2}\mathbf{b}'\mathbf{b})) \\
&= log((2\pi\sigma_z^2)^{-\frac{q}{2}}) - \frac{1}{2\sigma_z^2}\mathbf{b}'\mathbf{b} \\
&= -\frac{q}{2}log(2\pi\sigma_z^2) \quad - \quad \frac{1}{2\sigma_z^2}\mathbf{b}'\mathbf{b} \\
&= -\frac{1}{2}(q * log(2\pi\sigma_z^2) + \frac{1}{\sigma_z^2}\mathbf{b}'\mathbf{b})
\end{aligned}
$$

c)

Consider logistic mixed effects model

$y_i \sim Ber(\mu_i)$

$logit(\mu_i) = \beta_0 + b_h$

$b_h \sim N(0, \sigma_z^2)$

Formalize $f(y_i, b_h)$.

$f(y_i) = \mu_i^{y_i}(1-\mu_i)^{1-y_i}$

$$
\begin{aligned}
f(y_i, b_h) &= f(y_i \mid b_h) f(b_h) \\
&= \left[\frac{exp(\beta_0 + b_h)}{1 + exp(\beta_0 + b_h)}\right]^{y_i} \left[\frac{1}{1 + exp(\beta_0 + b_h)}\right]^{1-y_i} f(b_h) \\
&= \frac{exp(y_i(\beta_0 + b_h))}{1 + exp(\beta_0 + b_h)} f(b_h) \\
&= \frac{exp(y_i(\beta_0 + b_h))}{1 + exp(\beta_0 + b_h)} \frac{1}{\sqrt{2\pi\sigma_z^2}} exp(-\frac{1}{2\sigma_z^2} b_h^2) \\
&= \frac{1}{\sqrt{2\pi\sigma_z^2}} \frac{1}{1 + exp(\beta_0 + b_h)} exp(-\frac{1}{2\sigma_z^2} b_h^2 + y_i(\beta_0 + b_h))
\end{aligned}
$$