# 9.3 EXERCISES

**BMED6420**
Brani Vidakovic; Fall 2018
Consult the class slides, hints, and cited literature for the solution of exercise problems.

**1. IHGA Models, Which is the Best?** The following example is built In an experiment conducted in the 1980s (Hendriksen et al. 1984[1]), 572 elderly people living in a number of villages in Denmark were randomized, 287 to a control (C) group (who received standard care) and 285 to an experimental group (E) who received standard care plus IHGA: a kind of preventive assessment in which each person's medical and social needs were assessed and acted upon individually. The important outcome was the number of hospitalizations during the three-year life of the study.

Table 1: Distribution of number of hospitalizations in the IHGA study over a two-year period

| | Number of Hospitalizations | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | $n$ | Mean | Variance |
| Control | 138 | 77 | 46 | 12 | 8 | 4 | 0 | 2 | 287 | 0.944 | 1.54 |
| Treatment | 147 | 83 | 37 | 13 | 3 | 1 | 1 | 0 | 285 | 0.768 | 1.02 |

We will propose several models to assess the IHGA treatment using deviance measure. Which one is the best?

- **Treatment Effect Additive; Model Normal.**
  Program `geriatric0.odc`:

```
model
{
for (i in 1:n.C)
    {
    visits.C[i] ~ dnorm(mu.C, prec.C)
     }
for(j in 1:n.E)
     {
    visits.E[j] ~ dnorm(mu.E, prec.E)
     }
```

[1]Hendriksen, C., Lund, E., Stromgard, E. (1984). Consequences of assessment and intervention among elderly people: a three year randomized controlled trial. *British Medical Journal*, **289**, 1522–1524. Data also analysed from he Bayesian point of view by David Draper (UCSC) and his team.

```
mu.C ~ dnorm(0, 0.0001)
mu.E ~ dnorm(0, 0.0001)
prec.C ~ dgamma(0.0001, 0.0001)
prec.E ~ dgamma(0.0001, 0.0001)
effect <- mu.E - mu.C
var.C <- 1/prec.C
var.E <- 1/prec.E
}

#data

list( visits.C = c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                            ... lines skept
   1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2,
   2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
   2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
   2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4,
   4, 5, 5, 5, 5, 7, 7 ), n.C = 287,
       visits.E = c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                          ...  lines skept
    0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
    1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
    1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
    1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
    1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
    2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
    2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,
    4, 4, 4, 5, 6 ), n.E = 285 )

#INITS

list( mu.C=1, mu.E=1,  prec.C=1, prec.E = 1 )
```

What is deviance? What are shortcomings of this model? **Ans.** Hospital days are integers, we use normal distribution. Maybe Poisson is better?

- **Treatment Effect Additive; Model Poisson.**
  Program `geriatric1.odc`:

```
model{
for (i in 1:n.C)
    {
```

```
  visits.C[i] ~ dpois(lambda.C)
    }
 for(j in 1:n.E)
    {
  visits.E[j] ~ dpois(lambda.E)
    }
lambda.C ~ dgamma(0.0001, 0.0001)
lambda.E ~ dgamma(0.0001, 0.0001)
effect  <-  lambda.E - lambda.C
}
```

DATA and INITS as before

What is deviance? What are shortcomings of this model? **Ans.** Treatment effect should not be additive, rather multiplicative. Effect at increase of hospital days from 0 to 1 is not the same as the effect of increase from 5 to 6 days. How about multiplicative effect?

- **Treatment Effect Multiplicative; Model Poisson.**
  Program `geriatric2.odc`:

```
model
{
for (i in 1:n)
    {
      y[i] ~ dpois(lambda[i] )
      log( lambda[i]) <- gamma.0 + gamma.1 * x[i]
    }
 gamma.0 ~ dnorm(0, 0.0001)
 gamma.1 ~ dnorm(0, 0.0001)
 lambda.C <- exp(gamma.0)
 lambda.E <- exp(gamma.0 + gamma.1 )
 meffect  <- exp( gamma.1 )
 }

 #DATA
list( y = c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                  ... lines deleted
 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,
 4, 4, 4, 5, 6 ),
    x  =   c( 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                ... lines deleted
                1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
```

```
                1, 1, 1, 1, 1 ), n = 572 )

#inits
list( gamma.0 = 0.0, gamma.1 = 0.0 )
```

What is deviance of this model? What are shortcomings? **Ans.** Poisson model may not be elastic enough, if the mean of data is not close to its to variance, Poisson may not be adequate...unless we account for under/overdispersion. How about adding a random effect?

- **Treatment Effect Multiplicative; Model Poisson with Random Effect.**
  Program `geriatric3.odc`:

```
model
{
for (i in 1:n)
    {
   y[i] ~ dpois(lambda[i] )
   log( lambda[i]) <- gamma.0 + gamma.1 * x[i] + eps[i]
   eps[i] ~ dnorm(0, tau.eps)
   }
  # Why random effect eps? Overdispersion
  m <- mean(y[])  #Info>Node Info
  var <- pow(sd(y[]),2)  #Info>Node Info
 gamma.0 ~ dnorm(0, 0.0001)
 gamma.1 ~ dnorm(0, 0.0001)
 tau.eps ~ dgamma(0.001, 0.001)
 sig.eps <- 1/sqrt(tau.eps)
 lambda.C <- exp(gamma.0)
 lambda.E <- exp(gamma.0 + gamma.1 )
 meffect  <- exp( gamma.1 )
 }

 DATA as before.
 INITS
 list(gamma.0 = -0.058, gamma.1 = -0.21,
  tau.eps = 2.0)
```

What is deviance of this model? What are shortcomings? **Ans.** Variance in data exceeds the Mean, suggesting excess of zeros, in this case. How about adding a using zero-inflated Poisson? This model is not built in, we will need to use zero-trick.

- **Treatment Effect Multiplicative; Model Zero-Inflated Poisson.**
  Program `geriatric4.odc`:

```
model{
C<- 10000
for (i in 1:n) {
   zeros[i] <- 0
   zeros[i] ~ dpois(z.mean[i])
   z.mean[i] <-   -ll[i]+C
   ll[i] <-   log( p0[x[i]+1] * equals(y[i],0) + (1-p0[x[i]+1])*lf[i] )
   lf[i] <-  exp(-lambda[x[i]+1]+y[i]*log(lambda[x[i]+1]) -
               loggam(y[i]+1))
   }
   for (j in 1:2){
        p0[j] ~ dbeta(1,1)
        log(lambda[j]) <- gamma.0 + gamma.1 * equals(j,2)
        y.mean[j] <- (1-p0[j])*lambda[j]
        y.var[j]     <-  ( 1-p0[j] ) * ( lambda[j]+p0[j]*lambda[j]*lambda[j] )
        di[j]         <-  y.var[j]/y.mean[j]
        }
 gamma.0 ~ dnorm(0, 0.01)
 gamma.1 ~ dnorm(0, 0.01)
 meffect  <- exp( gamma.1 )
 Deviance <- -2*sum(ll[1:n])
 }
```

What is deviance of this model? What are shortcomings? **Ans.** Poisson model is often suboptimal to Negative Binomial for modeling purposes. After all, NB model has two parameters compared to single parameter Poisson. How about using Zero-Inflated Negative Binomial?

**• Treatment Effect Multiplicative; Model Zero-Inflated Negative Binomial with Random Effect.**

Program `geriatric5.odc`:

```
model{
C<- 10
for (i in 1:n) {
      #zero inflated NB via zero-trick
      zeros[i] <- 0
      zeros[i] ~ dpois(z.mean[i])
      z.mean[i] <-   -loglik[i]+C  #C ensures positive z.mean
      loglik[i] <-   log( p0[x[i]+1] * equals(y[i],0) +
                          (1-p0[x[i]+1])*nbpart[i] )
   nbpart[i] <-  exp(loggam( y[i]+r.ind[i] ) - loggam( r.ind[i] ) -
          loggam( y[i]+1 ) + r.ind[i]*log( p.ind[i] ) + y[i]*log( 1-p.ind[i] ))
   log(lambda[i]) <-  gamma.0 + gamma.1 * x[i] + eps[i]
```

```
   eps[i] ~ dnorm(0, tau.eps)
lfd[i] <-  loggam( y[i]+r.ind[i] ) - loggam( r.ind[i] ) -
loggam( y[i]+1 ) + r.ind[i]*log( p.ind[i] ) + y[i]*log( 1-p.ind[i] )
fd[i] <- exp( lfd[i] )
p.ind[i] <- r.ind[i]/( r.ind[i]+lambda.ind[i] )
r.ind[i] <- r[ x[i] + 1 ]
log(lambda.ind[i]) <- beta[1] + beta[2] * x[i]
   }
       for (j in 1:2){
             p0[j] ~ dbeta(1,1)
             lam[j] <- exp(gamma.0 + gamma.1 * equals(j,2) )
               y.mean[j] <- (1-p0[j])*lam[j]
               y.var[j]     <-  ( 1-p0[j] ) * ( lam[j]+p0[j]*lam[j]*lam[j] )
               di[j]         <-  y.var[j]/y.mean[j]
             }
 tau.eps ~ dgamma(0.001, 0.001)
 gamma.0 ~ dnorm(0, 0.01)
 gamma.1 ~ dnorm(0, 0.01)
 meffect  <- exp( gamma.1 )
 Deviance <- -2*sum(loglik[1:n])
 }
DATA (the same as for multiplicative effect)
INITS
list(gamma.0 = 1, gamma.1 = 1, tau.eps=1, p0=c(0.5, 0.5))
    #and generate the rest
```

Now real exercise: • **Zero-Truncated Poisson Regression for Number of Visits for Patients Who Checked to the Hospital at Least Once.**
*Hint:* Zero truncated Poisson has the same log-likelihood (up to additive constant) as the standard Poisson. In data, ignore 0's and adjust for the sample size.