SDS Page No. Date

Bayesian Computing

Name: Preksha Ashok Patel
Sapid: 60004210126
Branch:- Computer Engineering
Div: C2

Experiment no. 6

Aim:- To implement a Beta-Binomial Model for Over dispersion using cancer mortality data.

Theory:-

Overdispersion:-

This refers to a situation where ~~whe~~ the variability in the observed data is greater than what is expected from a given statistical model. In the context of a binomial distribution overdispersion occurs when the observed variance of the data exceeds the variance predicted by the assumed distribution. Overdispersion is common in real-world data, and failing to account for it may lead to underestimated uncertainties in statistical inferences.

Beta-Binomial Model:-

The beta-binomial model is a probabilistic model used to address overdispersion in binomially distributed data. It extends the standard binomial distribution by introduction additional variability through a beta distribution. In this model, each trial of the binomial distribution is associated with a random probability of success, drawn from a beta distribution. The beta-binomial distribution allows for a more flexible representation of variability in the data. The probability Mass function (PMF) of beta-binomial distribution is as follows:-

$$P(y_i/\theta_i) = \frac{{}^{n_i}C_{y_i} \; B(y_i+\alpha, \; n_i-y_i+\beta)}{B(\alpha, \beta)} \text{, where,}$$

P.T.O.

y_i is the number of successes in n_i trials, θ_i is the probability of success for the i-th trial, B() is the beta function, α & β are shape parameters

The prior distribution for shape parameters is : $\pi(\alpha, \beta) \propto 1$

The flat non-informative prior expresses a lack of prior knowledge about the shape parameters.

The posterior distribution for probability of success is :

$\pi(\theta_i / y_i) \propto P(y_i / \theta_i) \, \pi(\alpha, \beta)$

The posterior distribution incorporates the likelihood and prior, allowing for Bayesian inference. These equations define the Beta-Binomial model, which introduces flexibility through the beta distribution to account for overdispersion in binomial data.

Using Beta-Binomial Model for overdispersion :

1. Model specification : Define the beta-binomial model by incorporating a beta distribution to introduce additional variability in the standard binomial likelihood.
2. Prior Distribution : choose appropriate prior distributions for beta-binomial model parameters, leveraging Bayesian methods to incorporate prior Beliefs.
3. Posterior Inference : utilize Bayesian computational methods (eg MCMC) to estimate the posterior distribution, updating beliefs based on observed data.
4. Model Comparison : Compare the beta-binomial model with a standard binomial model to ~~ass~~ assess improvements in fit & quantify overdispersion.
5. Incorporating uncertainty : Leverage Bayesian Inference to naturally provide ~~leverage~~ estimates of parameter uncertainty, crucial for reliable prediction and understanding data ~~into~~ variability.

Conclusion : In this experiment, we implemented a beta-binomial model for overdispersion using cancer mortality data.

SAP ID.: 60004210126 Name: Preksha Ashok Patel

Batch: C2-1

Subject: Bayesian Computing Laboratory Semester: VII

Experiment No. 6

Aim:

Implement a Beta-Binomial Model for Over dispersion using Cancer mortality data.

Code:

Importing Libraries:

```
library(LearnBayes)
```

Contour plot of parameters η and K in the beta-binomial model:

```
mycontour(betabinexch0,

        c(.0001, .003, 1, 20000),
cancermortality,
xlab="eta", ylab="K")
```

Contour plot of transformed parameters logit(η) and log K in the beta-binomial model:

```
mycontour(betabinexch,

        c(-8, -4.5, 3, 16.5),
cancermortality,          xlab="logit
eta", ylab="log K")
```

Using 'laplace' for beta-binomial modelling:

```
fit <- laplace(betabinexch,

           c(-7, 6),
cancermortality)
```

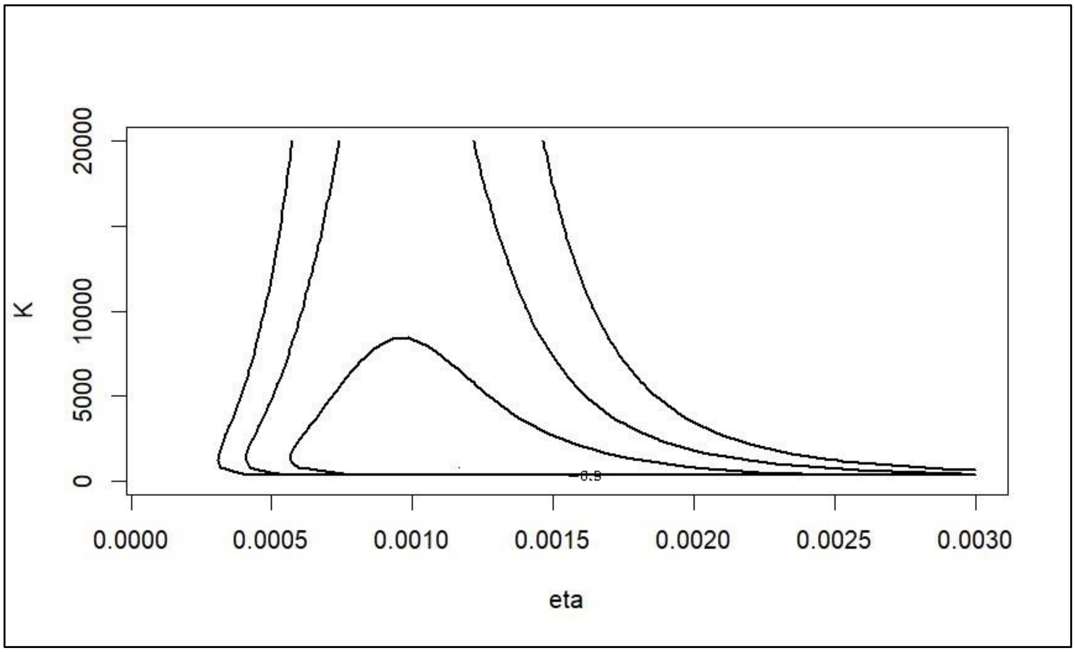
```
fit
```

Contour plot of normal approximation of logit(η) and log K in the beta-binomial model

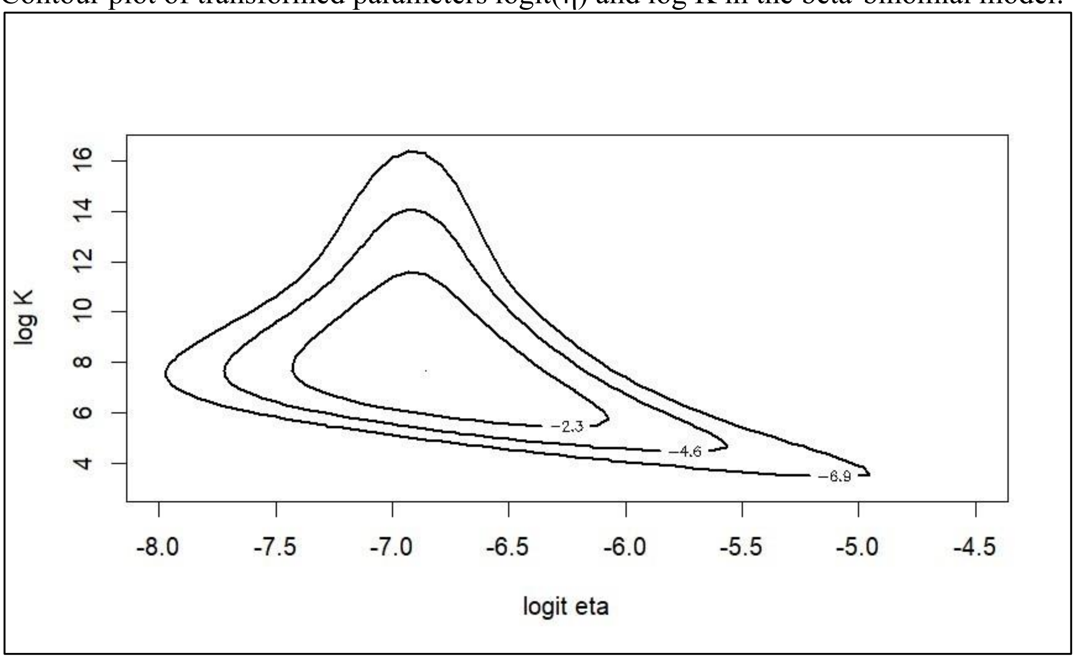
```
npar <- list(m=fit$mode, v=fit$var)

mycontour(lbinorm,

        c(-8, -4.5, 3, 16.5),

npar,         xlab="logit eta",

ylab="log K")


se <- sqrt(diag(fit$var))

fit$mode - 1.645 * se



fit$mode + 1.645 * se
```

Output:

Contour plot of parameters η and K in the beta-binomial model:



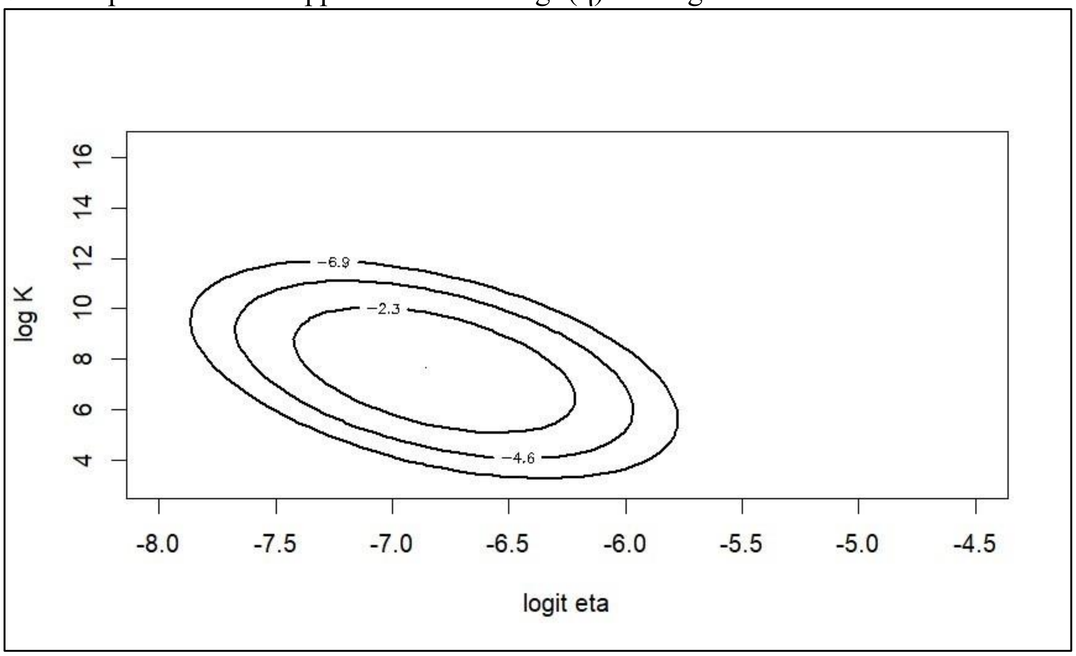
Contour plot of transformed parameters logit(η) and log K in the beta-binomial model:



Using ‘laplace’ for beta-binomial modelling:

```
$mode
[1] -6.819793  7.576111

$var
            [,1]       [,2]
[1,]  0.07896568 -0.1485087
[2,] -0.14850874  1.3483208

$int
[1] -570.7743

$converge
[1] TRUE
```

Contour plot of normal approximation of logit(η) and log K in the beta-binomial model




```
[1] -7.282052  5.665982
```

```
[1] -6.357535  9.486239
```