STAT40180 — Stochastic Models

Brendan Murphy

Week 2

Practical Inference

Practical Statistical Inference

- In this class, we will recap on maximum likelihood inference.
- The emphasis will be on using computational methods.
- We will consider the problems of:
  - Estimation
  - Interval Estimation
  - Testing

Time-to-Event: Kiama Blowhole

- Let's consider a time-to-event dataset.
- Spectacular eruptions of water are formed through a hole in the cliff at Kiama, about 120km south of Sydney, Australia.
- The times at which 64 successive eruptions occurred were observed.

Time-to-Event: Kiama Blowhole

- Let's consider a time-to-event dataset.
- Spectacular eruptions of water are formed through a hole in the cliff at Kiama, about 120km south of Sydney, Australia.
- The times at which 64 successive eruptions occurred were observed.

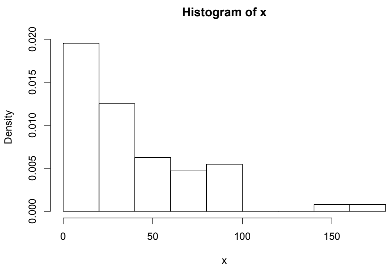
- The data are:

83	51	87	60	28	95	8	27	15	10	18	16	29	54	91	8
17	55	10	35	47	77	36	17	21	36	18	40	10	7	34	27
28	56	8	25	68	146	89	18	73	69	9	37	10	82	29	8
60	61	61	18	169	25	8	26	11	83	11	42	17	14	9	12

Time-to-Event: Model

- A potential model for the data is the exponential model

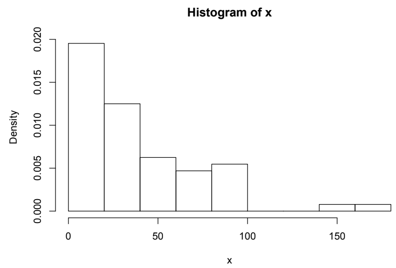
$$f(x) = \theta \exp(-\theta x), \text{ where } \theta > 0.$$



Time-to-Event: Model

- A potential model for the data is the exponential model

$$f(x) = \theta \exp(-\theta x), \text{ where } \theta > 0.$$



- The exponential can be easily fitted using maximum likelihood.

Maximum likelihood

- The likelihood is of the form

$$L(\theta) = \prod_{i=1}^{n} f(x_i)$$

Maximum likelihood

- The likelihood is of the form

$$L(\theta) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \theta \exp(-\theta x_i)$$

Maximum likelihood

- The likelihood is of the form

$$L(\theta) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \theta \exp(-\theta x_i) = \theta^n \exp\left(-\theta \sum_{i=1}^{n} x_i\right)$$

Maximum likelihood

- The likelihood is of the form

$$L(\theta) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \theta \exp(-\theta x_i) = \theta^n \exp\left(-\theta \sum_{i=1}^{n} x_i\right)$$

- Thus, the log-likelihood is

$$\ell(\theta) = n \log \theta - \theta \sum_{i=1}^{n} x_i.$$

- This is straightforward to maximize.
- We compute

$$S(\theta) = \frac{\partial}{\partial \theta} \ell(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

Maximum likelihood 2

- Solving

$$S(\theta) = 0$$

gives

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} x_i} = 1/\overline{x}.$$

- We check if the second derivative is negative when $\theta = \hat{\theta}$.
- Equivalently, we check if minus the second derivative is positive.

Maximum likelihood 2

- Solving

$$S(\theta) = 0$$

gives

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} x_i} = 1/\overline{x}.$$

- We check if the second derivative is negative when $\theta = \hat{\theta}$.
- Equivalently, we check if minus the second derivative is positive.

$$I(\theta) = -\frac{\partial^2}{\partial \theta^2} \ell(\theta) = -\frac{\partial}{\partial \theta} S(\theta) = \frac{n}{\theta^2}.$$

Maximum likelihood 2

- Solving
$$S(\theta) = 0$$
gives
$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} x_i} = 1/\overline{x}.$$
- We check if the second derivative is negative when $\theta = \hat{\theta}$.
- Equivalently, we check if minus the second derivative is positive.
$$I(\theta) = -\frac{\partial^2}{\partial \theta^2}\ell(\theta) = -\frac{\partial}{\partial \theta}S(\theta) = \frac{n}{\theta^2}.$$
- Clearly,
$$I(\hat{\theta}) > 0.$$

Maximum likelihood 3

- The estimator of θ is
$$\hat{\theta} = 1/\overline{x}.$$
- The approximate standard error is
$$SE(\hat{\theta}) = \sqrt{1/I(\hat{\theta})} = 1/\sqrt{I(\hat{\theta})}.$$

Maximum likelihood 3

- The estimator of θ is
$$\hat{\theta} = 1/\overline{x}.$$
- The approximate standard error is
$$SE(\hat{\theta}) = \sqrt{1/I(\hat{\theta})} = 1/\sqrt{I(\hat{\theta})}.$$
- Thus, an approximate 95% confidence interval would be
$$\hat{\theta} \pm \frac{1.96}{\sqrt{I(\hat{\theta})}}.$$

Maximum likelihood 3

- The estimator of θ is
$$\hat{\theta} = 1/\overline{x}.$$
- The approximate standard error is
$$SE(\hat{\theta}) = \sqrt{1/I(\hat{\theta})} = 1/\sqrt{I(\hat{\theta})}.$$
- Thus, an approximate 95% confidence interval would be
$$\hat{\theta} \pm \frac{1.96}{\sqrt{I(\hat{\theta})}}.$$
- For the Kiama blowhole data we get
$$\hat{\theta} \pm \frac{1.96}{\sqrt{I(\hat{\theta})}} = 0.025 \pm 0.006$$

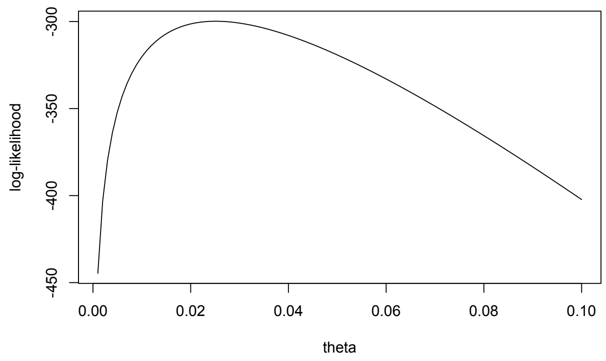
Maximum Likelihood: Code

- The code for doing the maximum likelihood estimation and inference.

```
x <- scan()
83  51  87  60  28  95   8  27  15  10  18  16  29  54  91   8
17  55  10  35  47  77  36  17  21  36  18  40  10   7  34  27
28  56   8  25  68 146  89  18  73  69   9  37  10  82  29   8
60  61  61  18 169  25   8  26  11  83  11  42  17  14   9  12

hist(x,probability=TRUE)

n <- length(x)
thetahat <- 1/mean(x)
se <- sqrt(thetahat^2/n)

UB <- thetahat+1.96*se
LB <- thetahat-1.96*se
```

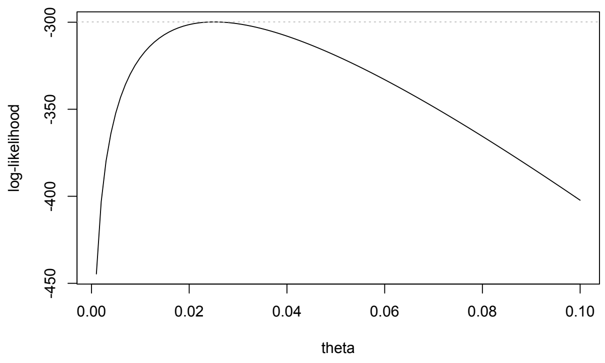

Maximum likelihood 4a

- This is how maximum likelihood works (in pictures).
- This is the likelihood function



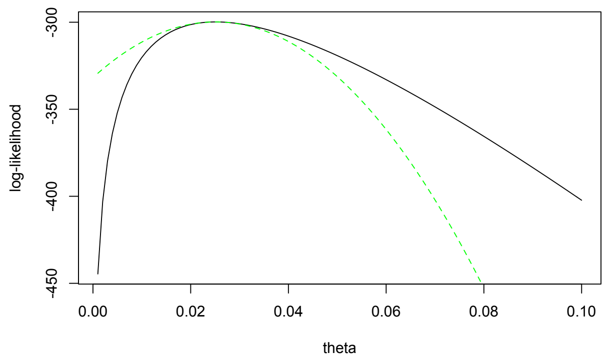
Maximum likelihood 4b

- We find where the function is flat ($S(\theta) = 0$).



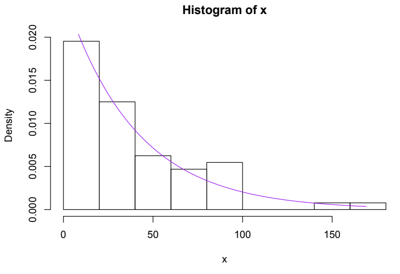
Maximum likelihood 4c

- We check that the function bends downwards at this point ($I(\hat{\theta}) > 0$)



Model Fit

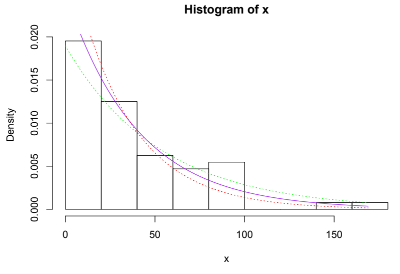
- We can informally compare the model fit to the data histogram.



- The model is fitting the data well.

Model Fit

- Adding the fits at the upper and lower end of the confidence interval helps assess fit.



Maximum Likelihood: Code

- The code for adding the fit to the histogram plot.

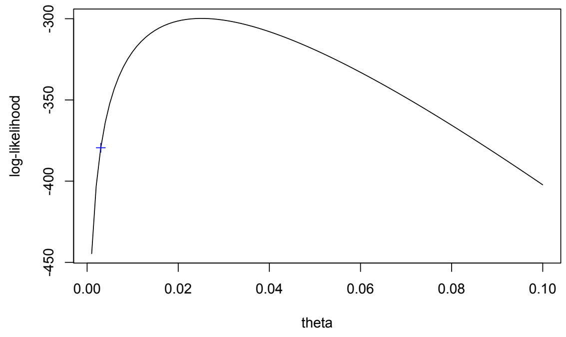
```
xseq <- seq(0,max(x),length=101)
points(xseq,dexp(xseq,thetahat),col="purple",type="l")

points(xseq,dexp(xseq,UB),col="red",type="l",lty=3)
points(xseq,dexp(xseq,LB),col="green",type="l",lty=3)
```


Maximum Likelihood: Computational Approach

- We could have taken a lazier approach to our model fitting.
- Once we had formed the likelihood, we could use a numerical optimizer to find the maximum likelihood estimator.
- For example, the Newton-Raphson algorithm can be used.
- I will demonstrate this using, pictures and then show how to do this in R.

Newton-Raphson 0a

- Suppose we have our log-likelihood function $\ell(\theta)$ and an initial guess at a value that maximizes it (call this θ_0).



- The guess wasn't very good!

Newton-Raphson 0b

- We can use a second order Taylor expansion to approximate the likelihood for θ near θ_0.

$$\ell(\theta) \approx \ell(\theta_0) + \ell'(\theta_0)(\theta - \theta_0) + \frac{1}{2}\ell''(\theta_0)(\theta - \theta_0)^2$$

Newton-Raphson 0b

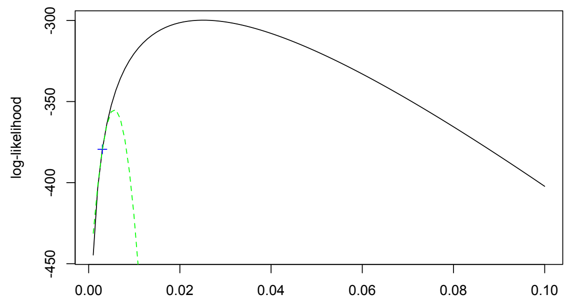
- We can use a second order Taylor expansion to approximate the likelihood for θ near θ_0.

$$\begin{aligned}\ell(\theta) &\approx \ell(\theta_0) + \ell'(\theta_0)(\theta - \theta_0) + \frac{1}{2}\ell''(\theta_0)(\theta - \theta_0)^2 \\ &= \ell(\theta_0) + S(\theta_0)(\theta - \theta_0) - \frac{1}{2}I(\theta_0)(\theta - \theta_0)^2\end{aligned}$$

Newton-Raphson 0b

- We can use a second order Taylor expansion to approximate the likelihood for θ near θ_0.

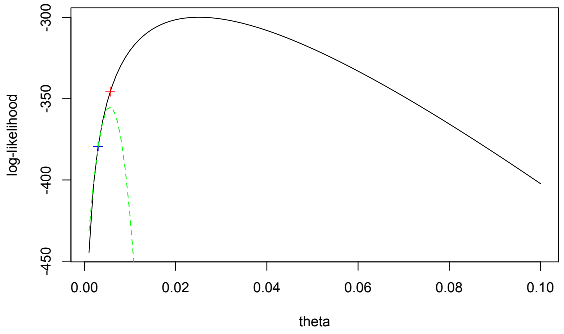
$$\begin{aligned}\ell(\theta) &\approx \ell(\theta_0) + \ell'(\theta_0)(\theta - \theta_0) + \frac{1}{2}\ell''(\theta_0)(\theta - \theta_0)^2 \\ &= \ell(\theta_0) + S(\theta_0)(\theta - \theta_0) - \frac{1}{2}I(\theta_0)(\theta - \theta_0)^2\end{aligned}$$



Newton-Raphson 0c

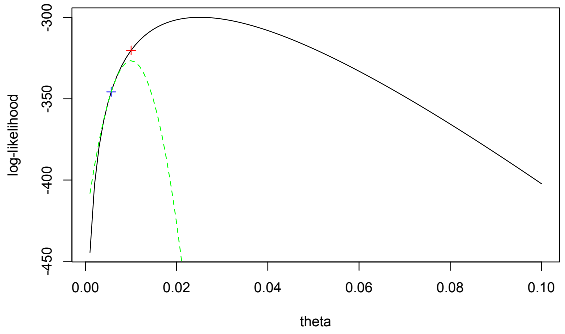
- We can find the θ that maximizes the approximation.

$$\theta_1 = \theta_0 + \frac{S(\theta_0)}{I(\theta_0)}$$



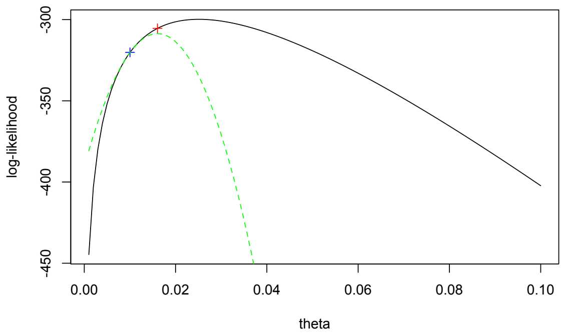
Newton-Raphson 1

- We now repeat the process of approximating the likelihood and then maximizing the approximation.



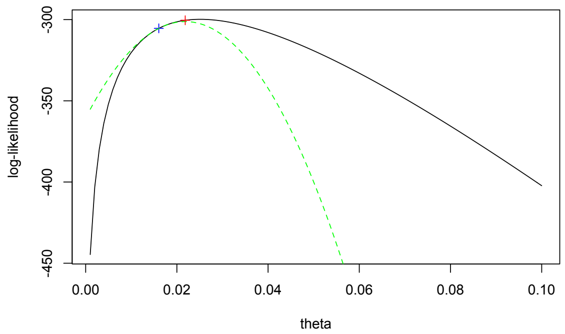
Newton-Raphson 2

- We now repeat the process of approximating the likelihood and then maximizing the approximation.



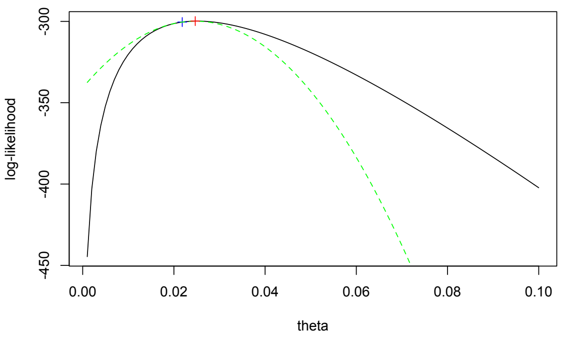
Newton-Raphson 3

- We now repeat the process of approximating the likelihood and then maximizing the approximation.



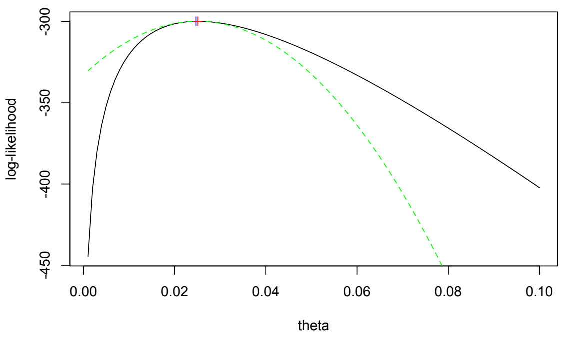
Newton-Raphson 4

- We now repeat the process of approximating the likelihood and then maximizing the approximation.



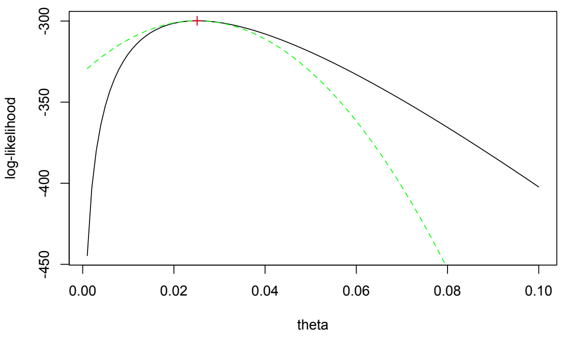
Newton-Raphson 5

- We now repeat the process of approximating the likelihood and then maximizing the approximation.



Newton-Raphson 6

- We now repeat the process of approximating the likelihood and then maximizing the approximation.



Newton-Raphson: Code

- We can use the `optim()` command in R to do this optimization.

```
loglik <- function(theta,x)
{
sum(dexp(x,theta,log=TRUE))
}
theta0 <- 0.03
fit <- optim(par=theta0,fn=loglik,method="BFGS",x=x,control=list(fnscale=-1),hessian=TRUE)
```

Newton-Raphson: Code

- We can use the optim() command in R to do this optimization.

```
loglik <- function(theta,x)
{
sum(dexp(x,theta,log=TRUE))
}
theta0 <- 0.03
fit <- optim(par=theta0,fn=loglik,method="BFGS",x=x,control=list(fnscale=-1),hessian=TRUE)
```

- We get the following output.

```
$par
[1] 0.02511322

$value
[1] -299.8127

$counts
function gradient
      51        5

$convergence
[1] 0

$message
NULL

$hessian
        [,1]
[1,] -101802
```


Newton-Raphson: Code

- The code can be more efficient if the derivative is also supplied.

```
score <- function(theta,x)
{
n<-length(x)
n/theta-sum(x)
}
fit <- optim(par=theta0,fn=loglik,gr=score,method="BFGS",x=x,control=list(fnscale=-1)
   ,hessian=TRUE)
```

- We get the following output.

```
$par
[1] 0.02510746

$value
[1] -299.8127

$counts
function gradient
      19        4

$convergence
[1] 0

$message
NULL

$hessian
          [,1]
[1,] -101686.6
```


STAT40180 — Stochastic Models

Brendan Murphy

Week 2

Multivariate Parameters

Multivariate Parameters

- In this class, we look at maximum likelihood inference for models with multivariate parameters.
- Again, the emphasis will be on using computational methods.

Time-to-Event: Kiama Blowhole

- Let's reconsider the Kiama blowhole data
- A more general model for the data is the Weibull model

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right], \text{ where } \alpha, \beta > 0.$$

- If $(\alpha, \beta) = (1, 1/\theta)$ then the Weibull model is the same as an exponential distribution.
- However, when $\alpha \neq 1$ it is has a different shape.
- Also,

$$\mathbb{E}(X) = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \text{ and } \mathbb{V}\text{ar}(X) = \beta^2\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma\left(1 + \frac{1}{\alpha}\right)^2\right]$$

So, method of moments would be very difficult for this model!

Maximum likelihood

- The likelihood is of the form

$$\begin{aligned} L(\alpha, \beta) &= \prod_{i=1}^{n} f(x_i) \\ &= \prod_{i=1}^{n} \frac{\alpha}{\beta}\left(\frac{x_i}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x_i}{\beta}\right)^{\alpha}\right] \\ &= \frac{\alpha^n}{\beta^n}\left(\frac{\prod_{i=1}^{n} x_i}{\beta^n}\right)^{\alpha-1} \exp\left[-\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\alpha}\right] \end{aligned}$$

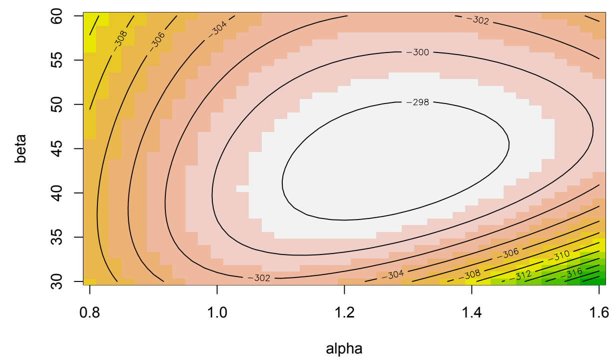
- The log-likelihood is cannot be maximized directly.

Maximum likelihood: Numerical

- We can maximize the likelihood using numerical methods.
- The method for maximization is very similar to the previous example, but we need to optimize with respect to α and β.
- For simplicity, we will let the unknown parameters be written as $\theta = (\theta_1, \theta_2) = (\alpha, \beta)$.
- We use the `optim()` command to optimize the log-likelihood with respect to the parameter θ.

Maximum likelihood: Likelihood

- We can produce a contour plot of the likelihood function to see how it varies with the value of $\theta = (\alpha, \beta)$.



Maximum Likelihood: Code

- The code for doing the maximum likelihood estimation.

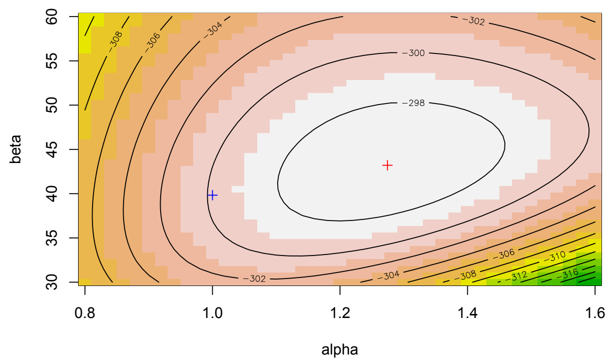
```
x <- scan()
83  51  87  60  28  95   8  27  15  10  18  16  29  54  91   8
17  55  10  35  47  77  36  17  21  36  18  40  10   7  34  27
28  56   8  25  68 146  89  18  73  69   9  37  10  82  29   8
60  61  61  18 169  25   8  26  11  83  11  42  17  14   9  12

loglik <- function(theta,x)
{
alpha <- theta[1]
beta <- theta[2]
sum(dweibull(x,alpha,beta,log=TRUE))
}

alpha0 <- 1
beta0 <- mean(x)
theta0 <- c(alpha0,beta0)
fit <- optim(par=theta0,loglik,method="BFGS",x=x,control=list(fnscale=-1),hessian=TRUE)
```

- We can see that

$$\hat{\theta} = (\hat{\alpha}, \hat{\beta}) = (1.27, 43.2).$$

Maximum likelihood: Likelihood

- The maximum likelihood estimate (red) and the exponential model fit (blue) are shown.



- The exponential model fit has log-likelihood value 2.91 lower than the maximum likelihood value.

Maximum Likelihood: Contour Code

- The code for doing the contour plot (with estimates) is below.

```
alphagrid <- seq(0.8,1.6,length=41)
betagrid <- seq(30,60,length=41)
thetagrid <- expand.grid(alphagrid,betagrid)
thetagrid <- as.matrix(thetagrid)

lvec<-rep(NA,41^2)
for (i in 1:nrow(thetagrid))
{
lvec[i] <- loglik(thetagrid[i,],x)
}

lmat <- matrix(lvec,41,41)

image(alphagrid,betagrid,lmat,col=terrain.colors(12),xlab="alpha",ylab="beta")
contour(alphagrid,betagrid,lmat,add=TRUE)

points(fit$par[1],fit$par[2],pch=3,col="red")
points(theta0[1],theta0[2],pch=3,col="blue")
```


Maximum likelihood: Inference

- The Hessian (second derivative) in this case is a matrix.
- Minus the Hessian is called the information matrix.
- The standard error of each parameters is given by

$$\mathrm{SE}(\hat{\theta}_1) = \sqrt{[I(\hat{\theta})^{-1}]_{11}} \text{ and } \mathrm{SE}(\hat{\theta}_2) = \sqrt{[I(\hat{\theta})^{-1}]_{22}}.$$

- That is, the square root of the diagonal of the inverse Information matrix gives the standard error for each parameter.

Maximum likelihood: Inference

- For the Weibull model we can use the following code to give the standard errors.

```
inf<- -fit$hessian
sqrt(diag(solve(inf)))
```

- We thus get,

$$SE(\hat{\alpha}) = 0.12 \text{ and } SE(\hat{\beta}) = 4.49.$$

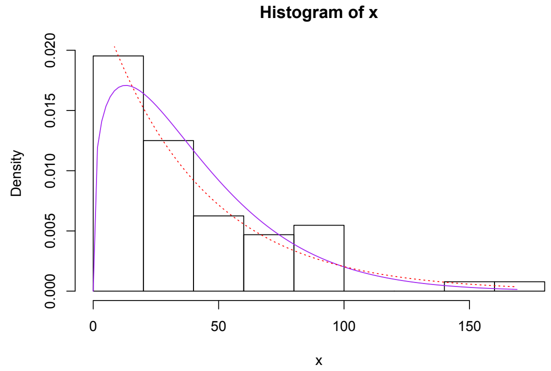
- Thus, approximate 95% confidence intervals for the parameters are:

$$\alpha : 1.274 \pm 0.236 \text{ and } \beta : 43.2 \pm 8.8.$$

- It is worth noting that 1 is not in the confidence interval for α.
- Thus, we have some evidence to support that $\alpha \neq 1$.

Model Fit

- We can informally compare the model fit (purple) to the data histogram.



- The exponential model is included (red) for comparison.

STAT40180 — Stochastic Models

Brendan Murphy

Week 2

Multiple Models

Multiple Models

- There are multiple plausible models for the Kiama blowhole data.
- For example:
  - Exponential
  - Weibull
  - Gamma
  - Log-normal
- We have already seen how to fit the first two models.
- The methods for the second two are similar.

Time-to-Event: Kiama Blowhole

- The gamma distribution also generalizes the exponential

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \text{ where } \alpha, \beta > 0.$$

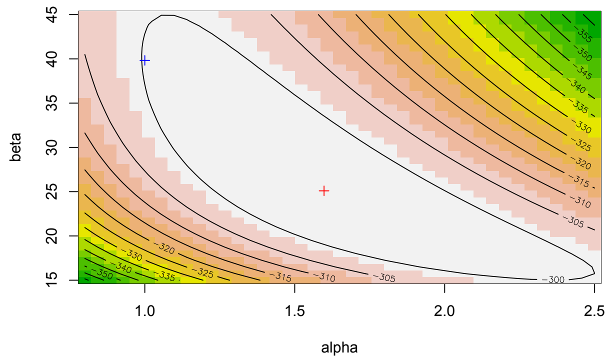
- If $(\alpha, \beta) = (1, 1/\theta)$ then the gamma model is the same as an exponential distribution.
- However, when $\alpha \neq 1$ it is has a different shape.
- Also,

$$\mathbb{E}(X) = \alpha\beta \text{ and } \mathbb{V}\text{ar}(X) = \alpha\beta^2$$

So, method of moments is straightforward for this model.

Maximum likelihood: Likelihood

- We can produce a contour plot of the likelihood function to see how it varies with the value of $\theta = (\alpha, \beta)$.



- The maximum is marked (red) and the exponential fit is also marked (blue).

Maximum Likelihood: Code

- The code for doing the maximum likelihood estimation.

```
loglik<-function(theta,x)
{
alpha<-theta[1]
beta<-theta[2]
sum(dgamma(x,shape=alpha,scale=beta,log=TRUE))
}

alpha0<-1
beta0<-mean(x)
theta0<-c(alpha0,beta0)
fit<-optim(par=theta0,loglik,method="BFGS",x=x,control=list(fnscale=-1),hessian=TRUE)
```

- We can see that

$$\hat{\theta} = (\hat{\alpha}, \hat{\beta}) = (1.60, 25.1)$$

and approximate 95% confidence intervals are:

$$\alpha : 1.60 \pm 0.52 \text{ and } \beta : 25.1 \pm 9.6.$$

Time-to-Event: Kiama Blowhole

- The log-normal distribution is another potential model

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right], \text{ where } \sigma > 0.$$

- This is equivalent to saying that $\log X \sim \text{Normal}(\mu, \sigma^2)$
- Also,

$$\mathbb{E}(X) = \exp(\mu + \sigma^2/2) \text{ and } \mathbb{V}\text{ar}(X) = \exp(2\mu + \sigma^2)\exp(\sigma^2 - 1)$$

In principle, method of moments can be done.

Maximum Likelihood: Code

- The code for doing the maximum likelihood estimation.

```
loglik<-function(theta,x)
{
mu<-theta[1]
sigma<-theta[2]
sum(dlnorm(x,meanlog=mu,sdlog=sigma,log=TRUE))
}

mu0<-mean(log(x))
sigma0<-sd(log(x))
theta0<-c(mu0,sigma0)
fit<-optim(par=theta0,loglik,method="BFGS",x=x,control=list(fnscale=-1),hessian=TRUE)
```

- We can see that

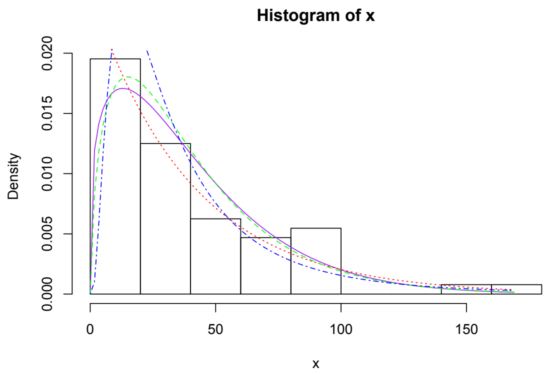
$$\hat{\theta} = (\hat{\mu}, \hat{\sigma}) = (3.35, 0.84)$$

and approximate 95% confidence intervals are:

$$\alpha : 3.35 \pm 0.11 \text{ and } \beta : 0.84 \pm 0.07.$$

Model Fit

- We can informally compare all of the model fits to the data histogram.



- Exponential=red, Weibull=purple, gamma=green,log-Normal=blue

Model Fit

- We can compare the models using the log-likelihood and the number of parameters.

Model	Log-likelihood	Parameters
Exponential	-299.8	1
Weibull	-296.9	2
Gamma	-295.9	2
Log-normal	-293.9	2

- We need to balance the quality of fit (log-likelihood) and model complexity (parameters).

Model Fit: Information Criteria

- Information criteria balance quality of fit ($\hat{\ell}$) and the number of parameters (p).

$$AIC = 2\hat{\ell} - 2p \text{ (Akaike Information Criterion)}$$

$$BIC = 2\hat{\ell} - log(n)p \text{ (Bayesian Information Criterion)}$$

Model	Log-likelihood	Parameters	AIC	BIC
Exponential	-299.8	1	-601.6	-603.8
Weibull	-296.9	2	-597.8	-602.1
Gamma	-295.9	2	-595.8	-600.1
Log-normal	-293.9	2	-591.8	-596.1

- The log-normal model has the highest AIC and BIC values[1].

[1] Some people define *AIC* and *BIC* in an equivalent but different manner

STAT40810 — Stochastic Models

Brendan Murphy

Week 2

Count Data

Count Data

- Suppose we have data that consist of an independent number of counts in the range $\{0, 1, 2, \ldots\}$.
- The standard model for such a situation is to model the data as Poisson, however alternatives exist:
  - Negative binomial
  - PoissonGamma
  - ConwayMaxwellPoisson (CMP or COM-Poisson) distribution
  - $\cdots$

Count Data:

- Data were collected on the number of houses in 1200 small (equal sized) areas in Japan.
- The number of houses per region were as follows:

Number	Frequency
0	584
1	398
2	168
3	35
4	9
5	4
7	1
9	1

- We want to find a suitable model for the number of houses per region.

Count Data: Poisson Model

- The Poisson model arises in the following situation.
- Suppose that $X_1, X_2, \ldots, X_n$ are independent count observations.
- If $X_i \sim \text{Poisson}(\lambda)$ then

$$f(x) = \mathbb{P}\{X_i = x\} = \frac{\lambda^x \exp(-\lambda)}{x!}, \text{ where } \lambda > 0.$$

- Under this model,

$$\mathbb{E}(X) = \lambda \text{ and } \mathbb{V}\text{ar}(X) = \lambda.$$

- So, the model cannot account for situations where

$$\mathbb{E}(X) \neq \mathbb{V}\text{ar}(X).$$

Count Data: Poisson-Gamma Model

- The Poisson-Gamma model arises in the following situation.
- Suppose that $X_1, X_2, \ldots, X_n$ are independent count observations.
- Assume that
$$X_i|\lambda_i \sim \text{Poisson}(\lambda_i)$$
and
$$\lambda_i \sim \text{Gamma}(\alpha, \beta).$$
That is, each subject has their own rate parameter in the Poisson and these are gamma distributed.
- Then,
$$X_i \sim \text{Poisson-Gamma}(\alpha, \beta).$$
- And,
$$f(x) = \frac{\Gamma(x+\alpha)\beta^x}{\Gamma(\alpha)(1+\beta)^{x+\alpha}\Gamma(x+1)}, \text{ where } \alpha, \beta > 0.$$
- Further,
$$\mathbb{E}(X) = \alpha\beta \text{ and } \mathbb{V}\text{ar}(X) = \alpha\beta + \alpha\beta^2.$$

Poisson-Gamma: Code

- The following code gives the log probability mass function of the Poisson-Gamma model.

```
dPoissonGamma <- function(x,alpha,beta,log=FALSE)
{
if (log)
{
lognumer <- lgamma(x+alpha)+x*log(beta)
logdenom <- lgamma(alpha)+(alpha+x)*log(1+beta)+lgamma(x+1)
res <- lognumer-logdenom
}else
{
numer <- gamma(x+alpha)*beta^x
denom <- gamma(alpha)*(1+beta)^(alpha+x)*gamma(x+1)
res <- numer/denom
}
res
}
```

- Setting `log=TRUE` gives the log probability mass function.

Modeling Question

(a) Find a method of moments estimate of the Poisson-Gamma model parameters.

(b) Fit the Poisson-Gamma model to the data using maximum likelihood.

(c) Fit the Poisson model to the data using maximum likelihood.

(d) Explain whether the Poisson or Poisson-Gamma model provide a better model for the data.

(e) Propose a method for assessing the fit of the models to the data and compare the method using your method.