Chapter 3: Exponential family responses and estimation

3.1 *Show that the geometric distribution, defined in exercise 2.3, is in the exponential family.*

$$\begin{aligned}\ln\{f(y)\} &= \ln\{(1-\pi)^y\pi\} \\ &= y\ln(1-\pi)+\ln\pi \\ &= \frac{y\theta-a(\theta)}{\phi}\ ,\end{aligned}$$

where $\theta=\ln(1-\pi)$, $a(\theta)=-\ln\pi=-\ln(1-e^\theta)$ and $\phi=1$.

3.2 (a) *Write down the likelihood equations for the estimation of μ and ν of the Gamma distribution.*

$$\begin{aligned}f(y) &= \frac{y^{-1}}{\Gamma(\nu)}\left(\frac{y\nu}{\mu}\right)^\nu \exp\left(-\frac{y\nu}{\mu}\right) \\ \ln f(y_i) &= \ell_i(\mu,\nu) \\ &= -\ln y_i-\ln\Gamma(\nu)+\nu\ln y_i+\nu\ln\nu-\nu\ln\mu-\frac{y_i\nu}{\mu}\end{aligned}$$

The likelihood is

$$\ell(\mu,\nu)=\sum_{i=1}^n\ell_i(\mu,\nu)=-\sum_i\ln y_i-n\ln\Gamma(\nu)+\nu\sum_i\ln y_i+n\nu\ln\nu-n\nu\ln\mu-\frac{\nu}{\mu}\sum_i y_i\ . \qquad (1)$$

Likelihood equation for μ:

$$\frac{\partial\ell(\mu,\nu)}{\partial\mu}=-\frac{n\nu}{\mu}+\frac{\nu}{\mu^2}\sum_i y_i=0 \qquad (2)$$

Likelihood equation for ν:

$$\frac{\partial\ell(\mu,\nu)}{\partial\nu}=-n\Psi(\nu)+\sum_i\ln y_i+n+n\ln\nu-n\ln\mu-\frac{1}{\mu}\sum_i y_i=0 \qquad (3)$$

where $\Psi(\nu)$ is the digamma function $\frac{\partial\ln\Gamma(\nu)}{\partial\nu}$.

(b) *Show that the mle of μ is $\hat\mu=\bar y$.*

Solving (2) for μ gives

$$\hat\mu=\frac{1}{n}\sum_i y_i=\bar y\ .$$

(c) *For R users:*

- *For claim size in the vehicle insurance data, write a program to estimate μ and ν.*

  The solution of (3) is not in closed form. We use the R function `optim` to maximise the log-likelihood directly. For maximisation of $\ell(\mu,\nu)$ with respect to ν, we substitute $\mu=\bar y$ in (1), and omit terms not involving ν :

$$
\begin{aligned}
\ell_\nu(\mu,\nu) &= -n\ln\Gamma(\nu)+\nu\sum_i \ln y_i + n\nu\ln\nu - n\nu\ln\bar{y} - \frac{\nu}{\bar{y}}\sum_i y_i \\
&= -n\ln\Gamma(\nu)+\nu\sum_i \ln y_i + n\nu(\ln(\nu/\bar{y}-1)
\end{aligned}
$$

```
 loglik.gamma<-function(v,y)
 {
 n<-length(y)
 mu<-mean(y)

 -(-n*lgamma(v)+v*sum(log(y))+n*v*(log(v/mu)-1))

    #### Return minus the log-likelihood,
    #### because optim will minimise this
 }

> attach(car)
> v<-optim(1,loglik.gamma,lower=0.01,y=claimcst0[claimcst0>0])$par
> v
[1] 0.75015
> mean(claimcst0[claimcst0>0])
[1] 2014.404
>
```

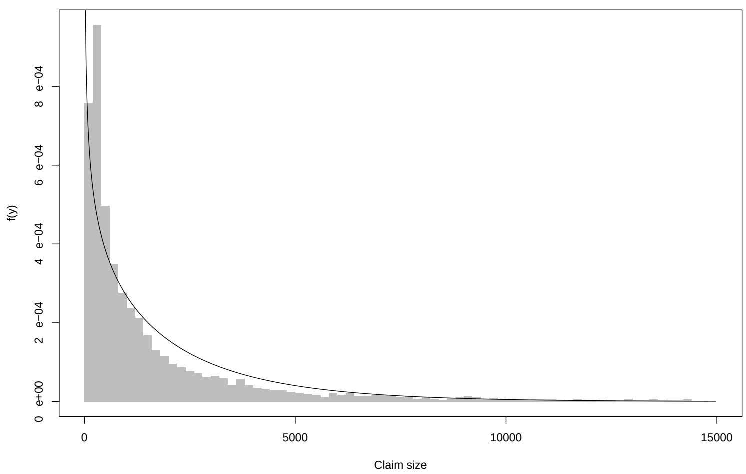
The solution is $\hat{\mu} = 2014.404, \hat{\nu} = 0.75015$. This may be confirmed (or obtained much more easily!) by fitting a null gamma model in SAS:

```
proc genmod data=act.car;
model claimcst0 = / dist=gamma link=id;
where claimcst0 > 0;
run;
```

```
                              The GENMOD Procedure

                               Model Information

                        Data Set                 ACT.CAR
                        Distribution               Gamma
                        Link Function           Identity
                        Dependent Variable     claimcst0


                    Number of Observations Read        4624
                    Number of Observations Used        4624

  .... [output omitted]

                          Analysis Of Parameter Estimates

                                 Standard     Wald 95% Confidence       Chi-
    Parameter      DF     Estimate       Error           Limits          Square    Pr > ChiSq

    Intercept       1     2014.404      34.2030    1947.367    2081.441    3468.69        <.0001
    Scale           1       0.7501       0.0134      0.7244      0.7768

NOTE: The scale parameter was estimated by maximum likelihood.
```

- Plot the fitted $G(\hat{\mu}, \hat{\nu})$ density, superimposed on a histogram of claim size.

```
gamplot<-function(y,max=15000,inc=200)
 {
 # y = data vector
 # max = x axis maximum
 # inc = histogram bin increment


 # Use hist function to get bin frequencies
 carhist<-hist(y,breaks=seq(0,max(y+inc),inc),include.lowest=T,freq=F,plot=F)

 ## Set up axes
  plot(1,0,xlim=c(0,max),ylim=c(0,max(carhist$density)),type="n",
    ylab="f(y)",xlab="Claim size")


 # Draw histogram
  for(i in 1:(length(carhist$breaks)-1))
    if(carhist$breaks[i+1]<max)
        rect(carhist$breaks[i],0,carhist$breaks[i+1],
        carhist$density[i],border=F,col="gray")

 # Compute Gamma MLEs
 v<-optim(1,loglik.gamma,lower=0.01,y=y)$par
 mu <- mean(y)

 # Gamma plot
 x<-seq(0.01,max,inc/10)
 f<-(1/(x*gamma(v)))*(((x*v/mu)^v))*exp(-x*v/mu)  # Gamma density
 lines(x,f)

 }

 gamplot(y=claimcst0[claimcst0>0],max=15000,inc=200)
```



3.3 (a) *Write down the likelihood equations for the estimation of μ and σ^2 of the Inverse*

Gaussian distribution.

$$\begin{aligned} f(y) &= \frac{1}{\sqrt{2\pi y^3}\sigma} \exp\left\{-\frac{1}{2y}\left(\frac{y-\mu}{\mu\sigma}\right)^2\right\} \\ \ln f(y_i) &= \ell_i(\mu,\nu) \\ &= -\frac{1}{2}\ln(2\pi y_i^3) - \ln\sigma - \frac{1}{2y_i}\left(\frac{y_i-\mu}{\mu\sigma}\right)^2 \\ &= -\frac{1}{2}\ln(2\pi y_i^3) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2y_i\sigma^2}\left(\frac{y_i-\mu}{\mu}\right)^2 . \end{aligned}$$

The likelihood is

$$\ell(\mu,\nu) = \sum_{i=1}^n \ell_i(\mu,\nu) = -\frac{1}{2}\sum_{i=1}^n \ln(2\pi y_i^3) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n \frac{1}{y_i}\left(\frac{y_i-\mu}{\mu}\right)^2 .$$

Likelihood equation for μ:

$$\frac{\partial\ell(\mu,\sigma^2)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^n \frac{y_i-\mu}{\mu^3} = 0 \qquad (4)$$

The trick is to differentiate with respect to σ^2, not σ. Likelihood equation for σ^2 :

$$\frac{\partial\ell(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n \frac{1}{y_i}\left(\frac{y_i-\mu}{\mu}\right)^2 = 0 \qquad (5)$$

(b) *Show that the mle of μ is $\hat\mu = \bar y$.*

Solving (4) for μ gives

$$\hat\mu = \frac{1}{n}\sum_i y_i = \bar y .$$

(c) *Show that the mle of σ^2 is $\hat\sigma^2 = n^{-1}\sum_i(1/y_i - 1/\bar y)$.*

Solving (5) for σ^2:

$$\begin{aligned} \hat\sigma^2 &= n^{-1}\sum_i \frac{1}{y_i}\left(\frac{y_i-\mu}{\mu}\right)^2 \\ &= n^{-1}\sum_i\left(\frac{y_i}{\mu^2} - \frac{2}{\mu} + \frac{1}{y_i}\right) \\ &= n^{-1}\sum_i \frac{1}{y_i} - \frac{1}{\bar y} \qquad \text{(substituting } \bar y \text{ for } \mu) . \end{aligned}$$

(d) *For R users:*

- *For claim size in the vehicle insurance data, estimate μ and σ^2.*

```
>  y <- claimcst0[claimcst0>0]
>  mu<-mean(y)
>  sigma2<-(1/length(y))*sum((1/y)-1/mean(y))
>  mu
[1] 2014.404
>  sigma2
[1] 0.001393199
```

We confirm this using SAS:

```
proc genmod data=act.car;
model claimcst0 = / dist=ig link=log ;
where claimcst0 > 0;
run;
```

```
                              The GENMOD Procedure

                               Model Information

                     Data Set                       ACT.CAR
                     Distribution          Inverse Gaussian
                     Link Function                      Log
                     Dependent Variable           claimcst0


                    Number of Observations Read        4624
                    Number of Observations Used        4624


                     Criteria For Assessing Goodness Of Fit

           Criterion                 DF           Value        Value/DF

           Deviance                4623          6.4422          0.0014
           Scaled Deviance         4623       4624.0000          1.0002
           Pearson Chi-Square      4623          7.1232          0.0015
           Scaled Pearson X2       4623       5112.8165          1.1060
           Log Likelihood                   -38591.8248


Algorithm converged.


                        Analysis Of Parameter Estimates

                               Standard     Wald 95% Confidence       Chi-
Parameter    DF    Estimate       Error           Limits             Square    Pr > ChiSq

Intercept     1      7.6081      0.0246      7.5598      7.6564    95369.2        <.0001
Scale         1      0.0373      0.0004      0.0366      0.0381

NOTE: The scale parameter was estimated by maximum likelihood.
```

The log link was used because the identity link resulted in an error message about convergence. We then get $\hat{\mu} = \mathrm{e}^{7.6081} = 2014.4$ and $\hat{\sigma}^2 = 0.0373^2 = 0.00139$. (The reported "Scale" is $\hat{\sigma}$.)

- *Plot the fitted* $IG(\hat{\mu}, \hat{\sigma}^2)$ *density, superimposed on a histogram of claim size.*

```
IGplot<-function(y,max=15000,inc=200)
 {
 # y = data vector
 # max = x axis maximum
 # inc = histogram bin increment


 # Use hist function to get bin frequencies
 carhist<-hist(y,breaks=seq(0,max(y+inc),inc),include.lowest=T,freq=F,plot=F)

 ## Set up axes
  plot(1,0,xlim=c(0,max),ylim=c(0,max(carhist$density)),type="n",
```

```
   ylab="f(y)",xlab="Claim size")


# Draw histogram
 for(i in 1:(length(carhist$breaks)-1))
   if(carhist$breaks[i+1]<max)
       rect(carhist$breaks[i],0,carhist$breaks[i+1],
       carhist$density[i],border=F,col="gray")

# Compute IG MLEs
mu<-mean(y)
sigma2<-(1/length(y))*sum((1/y)-1/mean(y))

## IG plot
x<-seq(0.01,max,inc/10)
sigma<-sqrt(sigma2)
f<-(1/(sqrt(2*pi*x^3)*sigma))*exp((-0.5/x)*(((x-mu)/(mu*sigma))^2))
lines(x,f)

}

IGplot(y=claimcst0[claimcst0>0],max=15000,inc=200)
```