Supplementary notes for new Part IIC course: Statistical Modelling, using the gamma distribution

P.M.E.Altham

January 4, 2005

We follow the notation of McCullagh and Nelder (1970), who define the density function of the gamma distribution with mean μ and shape parameter ν as

$$f(y|\mu,\nu) = \frac{1}{\Gamma(\nu)}(\nu y/\mu)^{\nu} e^{-\nu y/\mu}\ \frac{1}{y}$$

for $y > 0$.
You may check that this gives

$$E(Y) = \mu,\ var(Y) = (\mu^2)/\nu,$$

and the density is of standard glm form with $\phi = 1/\nu$, and canonical link $\eta = 1/\mu$.
We simulate from two gamma distributions below, and use the glm fitting procedure, with canonical link (ie the inverse). See if you can work out what's going on.

```
> library(MASS)
> y1 = rgamma(100, shape=5, rate=0.1)
> y2 = rgamma(50, shape=5, rate= 1.0)
> par(mfrow=c(2,1))
> truehist(y1) ; truehist(y2) # graphs not shown here
summary(y1); summary(y2)
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
4.817  30.770  43.770  48.360  59.730 114.300
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1.173   3.000   4.716   4.694   6.083  10.410

> x =c(rep("a", times=100), rep("b", times=50))
> is.factor(x) ; x = factor(x)
> y = c(y1,y2)
> plot(x,y)   # graphs not shown here
> first.glm = glm(y~x, Gamma) # nb, do not use "gamma"
> summary(first.glm)
Call:
glm(formula = y ~ x, family = Gamma)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
     -1.21730  -0.33643  -0.09652   0.25573   1.10905

 Coefficients:
             Estimate Std. Error t value Pr(>|t|)
 (Intercept) 0.0209790  0.0009124   22.99   <2e-16 ***
```

```
xb           0.1720752  0.0119088   14.45   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.1891440)

Null deviance: 144.609  on 149  degrees of freedom
Residual deviance:  28.430  on 148  degrees of freedom
  AIC: 1095.8
> dev = residuals(first.glm, type="deviance")
>summary(dev) ; sum(dev^2)
```

This fits $1/\mu_i = \alpha$ for the first 100 observations, and $1/\mu_i = \alpha + \beta$ for the remaining 50 observations. What is $1/\hat{\alpha}$?
What is $1/(\hat{\alpha} + \hat{\beta})$?
Note that the estimate given for ν is the reciprocal of the dispersion parameter ϕ, and this dispersion parameter is estimated by

$$X^2/(n-p)$$

where n is the number of observations, and p is the number of parameters in the linear model (here $p = 2$) and

$$X^2 = \Sigma[(y_i - \hat{\mu}_i)/\hat{\mu}_i]^2$$

Thus we find for this example that $\hat{\nu} = 5.287$. This is actually a 'moments' estimator rather than the mle: as an exercise you can write down the equation for the maximum likelihood extimator. You will find that this gives an equation involving the function $\Gamma'(\nu)/\Gamma(\nu)$ (the digamma function), and there is no closed-form solution to this equation.
I must admit, I had difficulty working out where the AIC came from. It is, I believe, minus twice the maximised log-likelihood $+2 \times 3$, since we were fitting 3 parameters. Try

```
> nu <- 5.287 # your simulation may mean you have a different estimator here
> fv <- first.glm$fitted.value
> term= -lgamma(nu) + nu*log(nu * y/fv) - (nu*y/fv) - log(y)
> sum(term)
  -544.9114
```

and I trust you will see what I mean.

Reference
P.McCullagh and J.A.Nelder *Generalized Linear Models* Chapman and Hall (1990).