Homework 5

Sarah Adilijiang

Problem 2

(a)

For $Y \sim Gamma(\alpha, \lambda)$, we have $f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}$ $(y \geq 0)$

Thus:

$$E(Y) = \int_0^\infty y f_Y(y) dy = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} y^\alpha e^{-\lambda y} dy = \frac{\alpha}{\lambda} \int_0^\infty \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} y^{(\alpha+1)-1} e^{-\lambda y} dy = \frac{\alpha}{\lambda}$$

$$E(Y^2) = \int_0^\infty y^2 f_Y(y) dy = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha+1} e^{-\lambda y} dy = \frac{\alpha(\alpha+1)}{\lambda^2} \int_0^\infty \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} y^{(\alpha+2)-1} e^{-\lambda y} dy = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\Rightarrow \quad Var(Y) = E(Y^2) - (E(Y))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}$$

Plug in $\alpha = \nu$, $\lambda = \nu/\mu$, we get: $E(Y) = \mu$, $Var(Y) = \mu^2/\nu$

So $\sigma^2 = Var(Y)/(E(Y))^2 = (\mu^2/\nu)/\mu^2 = 1/\nu$

(b)

From question (a), we get: $E(Y) = \mu$, $Var(Y) = E(Y-\mu)^2 = \mu^2/\nu$, $\sigma^2 = 1/\nu$

Set $g(Y) = \log Y$, where $g'(Y) = \frac{1}{Y}$ and $g''(Y) = -\frac{1}{Y^2}$

From Taylor expansion, we have:

$$g(Y) = g(\mu) + \frac{g'(\mu)}{1!}(Y-\mu) + \frac{g''(\mu)}{2!}(Y-\mu)^2 + ... + \frac{g^{(k)}(\mu)}{k!}(Y-\mu)^k + ...$$

$$i.e. \quad \log Y = \log \mu + \frac{1}{\mu}(Y-\mu) - \frac{1}{2\mu^2}(Y-\mu)^2 + O((Y-\mu)^3)$$

$$\Rightarrow \quad E(\log Y) \approx \log \mu + \frac{1}{\mu}E(Y-\mu) - \frac{1}{2\mu^2}E(Y-\mu)^2 = \log \mu - \frac{1}{2\mu^2}\mu^2/\nu = \log \mu - \frac{1}{2\nu} = \log \mu - \sigma^2/2$$

$$Var(\log Y) \approx Var(\log \mu + \frac{1}{\mu}(Y-\mu)) = \frac{Var(Y)}{\mu^2} = \frac{\mu^2/\nu}{\mu^2} = 1/\nu = \sigma^2$$

(c)

For $Y \sim Gamma(\alpha, \lambda)$, we have $f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}$ $(y \geq 0)$

Set $X = \log Y$, i.e. $Y = e^X$, so we have: $E(X) = M'(t)|_{t=0}$, $E(X^2) = M''(t)|_{t=0}$

Thus:

$$M(t) = E(e^{xt}) = E(y^t) = \int_0^\infty y^t f_Y(y) dy = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha+t-1} e^{-\lambda y} dy = \frac{\Gamma(\alpha+t)}{\Gamma(\alpha)\lambda^t} \int_0^\infty \frac{\lambda^{\alpha+t}}{\Gamma(\alpha+t)} y^{(\alpha+t)-1} e^{-\lambda y} dy = \frac{\Gamma(\alpha+t)}{\Gamma(\alpha)\lambda^t}$$

$$\Rightarrow \quad M'(t) = \frac{\Gamma'(\alpha+t)\Gamma(\alpha)\lambda^t - \Gamma(\alpha+t)\Gamma(\alpha)\lambda^t \log \lambda}{\Gamma(\alpha)^2 \lambda^{2t}} = \frac{\Gamma'(\alpha+t) - \Gamma(\alpha+t)\log \lambda}{\Gamma(\alpha)\lambda^t}$$

$$E(\log Y) = E(X) = M'(t)|_{t=0} = \frac{\Gamma'(\alpha) - \Gamma(\alpha)\log \lambda}{\Gamma(\alpha)} = \psi(\alpha) - \log \lambda$$

1

Then:
$$\Rightarrow \quad M''(t) = \frac{\Gamma''(\alpha + t) - 2\Gamma'(\alpha + t)\log\lambda + \Gamma(\alpha + t)(\log\lambda)^2}{\Gamma(\alpha)\,\lambda^t}$$

$$E(X^2) = M''(t)|_{t=0} = \frac{\Gamma''(\alpha) - 2\Gamma'(\alpha)\log\lambda + \Gamma(\alpha)(\log\lambda)^2}{\Gamma(\alpha)} = \frac{\Gamma''(\alpha)}{\Gamma(\alpha)} - 2\psi(\alpha)\log\lambda + (\log\lambda)^2$$

Since $\psi'(x) = \frac{\Gamma''(x)\Gamma(x) - \Gamma'(x)^2}{\Gamma(x)^2} = \frac{\Gamma''(x)}{\Gamma(x)} - \psi(x)^2$, so we have:

$$E(X^2) = \psi'(\alpha) + \psi(\alpha)^2 - 2\psi(\alpha)\log\lambda + (\log\lambda)^2 = \psi'(\alpha) + (\psi(\alpha) - \log\lambda)^2$$

$$\Rightarrow \quad Var(\log Y) = Var(X) = E(X^2) - (E(X))^2 = \psi'(\alpha)$$

Plug in $\alpha = \nu$, $\lambda = \nu/\mu$, $\sigma^2 = 1/\nu$, we get:

$$E(\log Y) = \psi(\nu) - \log(\nu/\mu) = \log\mu + \psi(\nu) - \log\nu$$

$$Var(\log Y) = \psi'(\nu)$$

(d)

Since independent observations $Y_i \sim G(\mu_i, \nu)$ for $i = 1, .., n$, so the likelihood function is:

$$L(\beta) = \prod_{i=1}^{n} f_Y(y_i) = \prod_{i=1}^{n} \frac{\nu^\nu}{\Gamma(\nu)\,\mu_i^\nu} y_i^{\nu-1} \exp(-\frac{\nu}{\mu_i} y_i)$$

Thus the log-likelihood function is:

$$l(\beta) = \log L(\beta) = \sum_{i=1}^{n} -\frac{\nu}{\mu_i} y_i - \nu\log\mu_i + \log\frac{\nu^\nu y_i^{\nu-1}}{\Gamma(\nu)}$$

Since we are using log-link: $\log\mu_i = \eta_i = \alpha + x_i^t\beta$, so $\mu_i = e^{\eta_i}$ and $\frac{\partial\mu_i}{\partial\beta} = \frac{\partial e^{\eta_i}}{\partial\beta} = e^{\eta_i}x_i = \mu_i x_i$

So the score function is:

$$S(\beta) = \frac{\partial l(\beta)}{\partial\beta} = \sum_{i=1}^{n} \frac{\nu}{\mu_i^2}\frac{\partial\mu_i}{\partial\beta} y_i - \frac{\nu}{\mu_i}\frac{\partial\mu_i}{\partial\beta} = \sum_{i=1}^{n} \frac{\nu}{\mu_i^2}\mu_i x_i y_i - \frac{\nu}{\mu_i}\mu_i x_i = \sum_{i=1}^{n} \frac{\nu}{\mu_i}x_i y_i - \nu x_i = \sum_{i=1}^{n} \frac{y_i - \mu_i}{\mu_i}\nu x_i$$

Therefore, the score function at the estimated $\hat\beta$ is: $S(\hat\beta) = \sum_{i=1}^{n} \frac{y_i - \hat\mu_i}{\hat\mu_i}\nu x_i$

Then the information matrix J is:

$$J = -E[\frac{\partial^2 l(\beta)}{\partial\beta^2}] = E[(\frac{\partial l(\beta)}{\partial\beta})^2] = E[S(\beta)^2] = E[S(\beta)S(\beta)^t]$$

$$= \sum_{i=1}^{n} \frac{E(y_i - \mu_i)^2}{\mu_i^2}\nu^2 x_i x_i^t = \sum_{i=1}^{n} \frac{Var(y_i)}{\mu_i^2}\nu^2 x_i x_i^t = \sum_{i=1}^{n} \frac{\mu_i^2/\nu}{\mu_i^2}\nu^2 x_i x_i^t = \sum_{i=1}^{n} \nu x_i x_i^t = \nu X^T X$$

where X is the design matrix without intercept terms.

Therefore, the information matrix at the estimated $\hat\beta$ is: $J_{\hat\beta} = \sum_{i=1}^{n} \nu x_i x_i^t = \nu X^T X$, where X is the design matrix without intercept terms.

Then, since $\hat\beta \sim N(\beta, J_\beta^{-1})$, we can get the covariance matrix of $\hat\beta$ is: $\Sigma_{\hat\beta} = J_{\hat\beta}^{-1} = (\nu X^T X)^{-1} = \frac{1}{\nu}(X^T X)^{-1}$.

Plug in the coefficient of variation $\sigma^2 = 1/\nu$ we get that the covariance matrix of $\hat\beta$ is: $\Sigma_{\hat\beta} = J_{\hat\beta}^{-1} = \sigma^2(X^T X)^{-1}$

(e)

Now we fit the linear model: $E(\log Y_i) = \alpha^* + x_i^t \beta^*$, i.e. $\log Y_i = \alpha^* + x_i^t \beta^* + \epsilon_i^*$, where $i.i.d.$ $\epsilon_i^* \sim N(0, \sigma^{*2})$

The closed form solution is: $\hat{\beta}^* = (X^T X)^{-1} X^T Z$, where $Z = \log Y$, and X is the design matrix without intercept terms.

We have: $\hat{\beta}^* \sim N(\beta, \sigma^{*2}(X^T X)^{-1})$, so the the covariance matrix of $\hat{\beta}^*$ is: $\Sigma_{\hat{\beta}^*} = \sigma^{*2}(X^T X)^{-1}$

From question (c), we get that the exact value of σ^{*2} is: $\sigma^{*2} = Var(\log Y) = \psi'(\nu)$. However, from question (b), when we're using the variance stabilizing transformation, we have the approximate value of σ^{*2}, which is: $\sigma^{*2} = Var(\log Y) \approx \sigma^2 = 1/\nu$

To sum up, the two covariance matrices have the same form:

(1) $\Sigma_{\hat{\beta}} = J_{\hat{\beta}}^{-1} = \sigma^2(X^T X)^{-1} = \frac{1}{\nu}(X^T X)^{-1}$, where σ^2 is the coefficient of variation in the $Gamma(\alpha = \nu, \lambda = \nu/\mu)$ distribution model, and $\sigma^2 = 1/\nu$

(2) $\Sigma_{\hat{\beta}^*} = \sigma^{*2}(X^T X)^{-1} = \psi'(\nu)(X^T X)^{-1}$, where σ^{*2} is the variance of errors in the linear model.

Though their exact values are not equal, for large n or small σ^* (or σ), approximately we have: $\sigma^{*2} \approx \sigma^2$, i.e. $\Sigma_{\hat{\beta}} \approx \Sigma_{\hat{\beta}^*}$, thus the two covariance matrices will be very similar.

(f)

(i)

```
load("fly.Rdata")

# relationship between du.sd & duration
m = lm(du.sd~duration, fly)
summary(m)
```

```
##
## Call:
## lm(formula = du.sd ~ duration, data = fly)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.056821 -0.026922 -0.015659  0.000101  0.246001
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0181440  0.0296901  -0.611  0.54768
## duration     0.0034144  0.0009212   3.707  0.00131 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06583 on 21 degrees of freedom
## Multiple R-squared:  0.3955, Adjusted R-squared:  0.3667
## F-statistic: 13.74 on 1 and 21 DF,  p-value: 0.001308
```

```
# plot du.sd ~ duration
plot(du.sd~duration, fly, xlab="mean of duration", ylab="standard deviation of duration")
x = seq(10,70,0.001)
lines(x, predict(m, data.frame(duration=x)), lty=2)
```
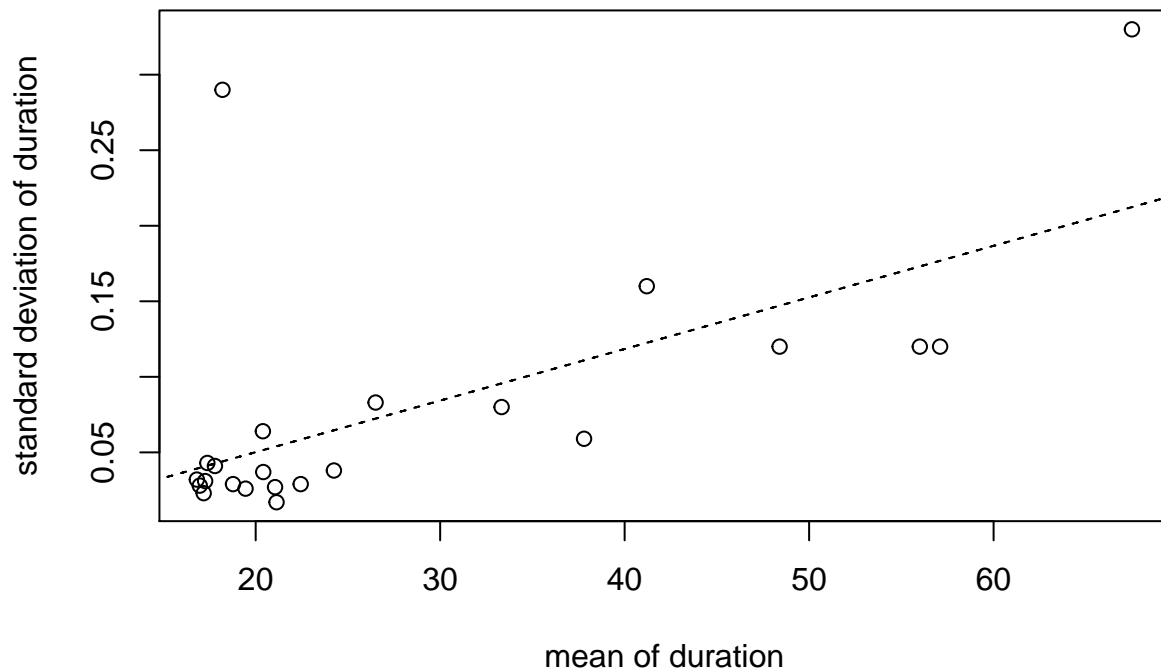
Answer:

From question (a), we get that for $Y \sim Gamma(\alpha = \nu, \lambda = \nu/\mu)$ distribution, we have: $\frac{SD(Y)}{E(Y)} = \frac{\sqrt{Var(Y)}}{E(Y)} = \sigma = 1/\sqrt{\nu}$, where σ^2 is the coefficient of variation. So there should be a linear relationship between $SD(Y)$ and $E(Y)$.

From the summary for our data, we can see that the standard deviation of the duration, i.e. $SD(Y)$, does have a significant linear relationship with the mean of the duration, i.e. $E(Y)$. This fitted linear relationship is also shown in the plots.

Also, the response "duration" is a positive value. Therefore, a Gamma GLM would make sense for this data.
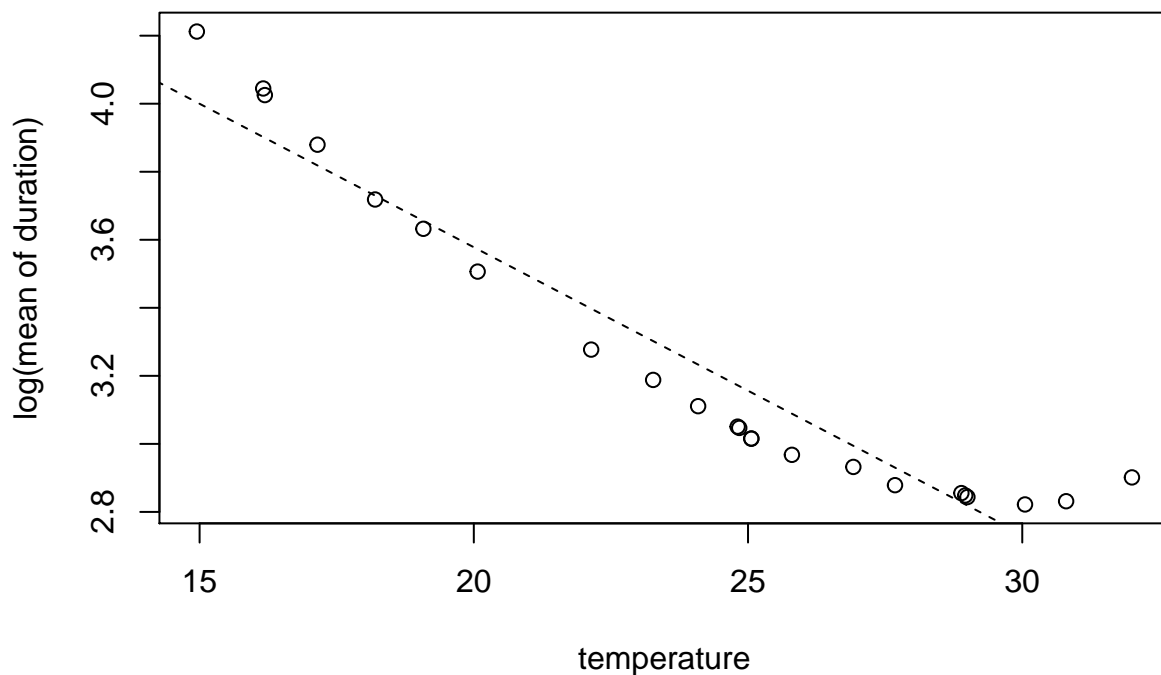
(ii)

```
# relationship of log(duration) ~ temp
m1 = lm(log(duration)~temp, fly)
summary(m1)

##
## Call:
## lm(formula = log(duration) ~ temp, data = fly)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -0.13617 -0.12067 -0.02337  0.07636  0.33565
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
```

4

```
## (Intercept)   5.265226    0.134956    39.01   < 2e-16 ***
## temp          -0.084358    0.005512   -15.30 7.31e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1325 on 21 degrees of freedom
## Multiple R-squared:  0.9177, Adjusted R-squared:  0.9138
## F-statistic: 234.2 on 1 and 21 DF,  p-value: 7.308e-13
```

```r
# plot log(duration) ~ temp
plot(log(duration)~temp, fly, xlab="temperature", ylab="log(mean of duration)")
x = seq(10,40,0.001)
lines(x, predict(m1, data.frame(temp=x)), lty=2)
```



```r
# relationship of log(duration) ~ 1/temp
m2 = lm(log(duration)~I(1/temp), fly)
summary(m2)
```

```
##
## Call:
## lm(formula = log(duration) ~ I(1/temp), data = fly)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.06161 -0.03606 -0.00382  0.02500  0.19046
##
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.39046    0.05026   27.66   <2e-16 ***
## I(1/temp)   42.25620    1.11526   37.89   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05546 on 21 degrees of freedom
## Multiple R-squared:  0.9856, Adjusted R-squared:  0.9849
## F-statistic:  1436 on 1 and 21 DF,  p-value: < 2.2e-16
```

```r
# plot log(duration) ~ 1/temp
plot(log(duration)~I(1/temp), fly, xlab="1/temperature", ylab="log(mean of duration)")
x = seq(10,40,0.001)
lines(1/x, predict(m2, data.frame(temp=x)), lty=2)
```
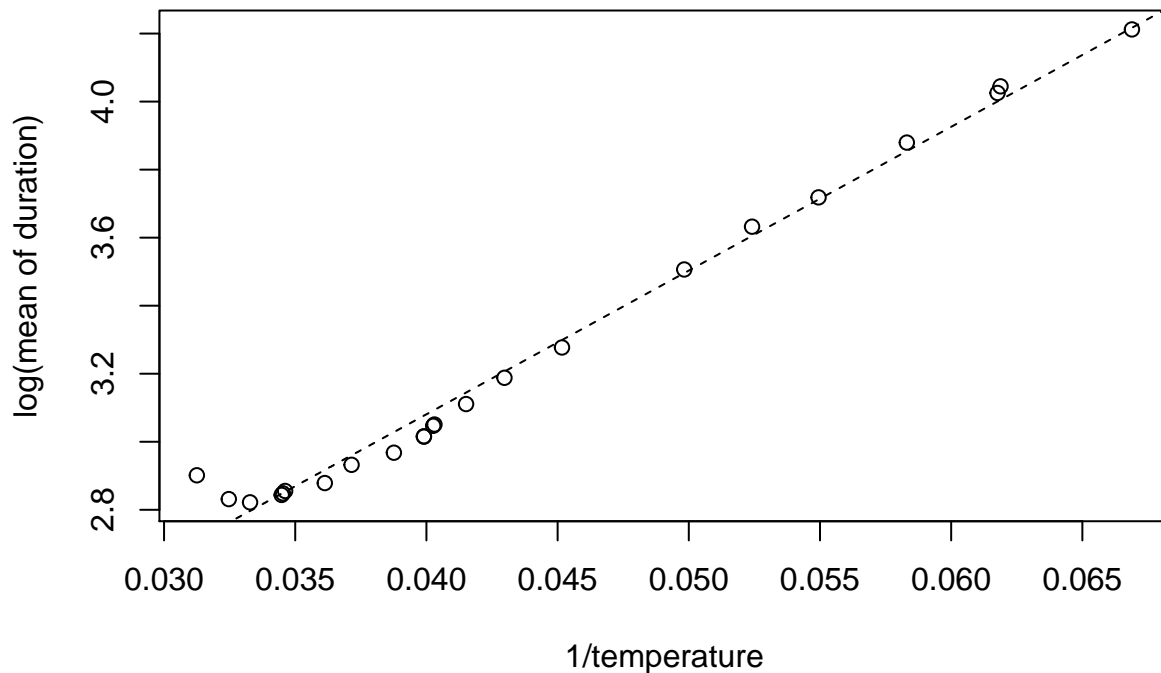

Answer:

The plot of "log(duration) ~ temp" presents a hyperbola function like curve, i.e. y~k/x, indicating a liear relationship of "log(duration) ~ 1/temp".

And I also tried to fit a linear line for both relationships. We can see that the linear relationship between "log(duration)" and "temp" is much less significant than the linear relationship between "log(duration)" and "1/temp". These fitted lines are also shown in the plots. It's obvious that the points lie much better in the line for "log(duration) ~ 1/temp".

Therefore, we should try to fit "duration ~ 1/temp" using log-link in the Gaussian GLM model.

(iii)

The GLM model for $Y = duration \sim Gamma(\alpha = \nu, \ \lambda = \nu/\mu)$ response here is:
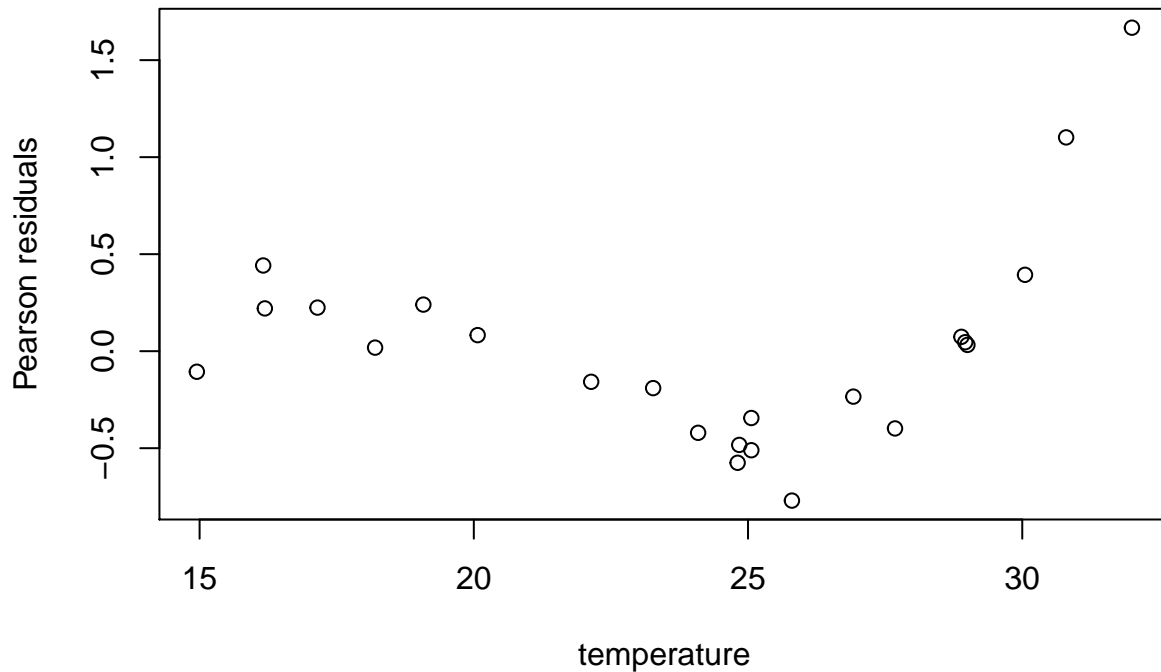
likelihood: $f(y_i) = \frac{\nu^\nu}{\Gamma(\nu)\,\mu_i^\nu} y_i^{\nu-1} \exp(-\frac{\nu}{\mu_i} y_i), \ y_i \geq 0$

linear predictor: $\eta_i = \beta_0 + \beta_1 \frac{1}{temp_i} + \epsilon_i$

link function (log-link): $\eta_i = \log\mu_i$, where $\mu_i = E(y_i)$

```r
# fit Gamma GLM
mod1 = glm(duration~I(1/temp), fly, family=Gamma(link=log), weights=batch)

# plot Pearson residuals ~ temp
plot(residuals(mod1,"pearson")~temp, fly,
     xlab="temperature", ylab="Pearson residuals")
```
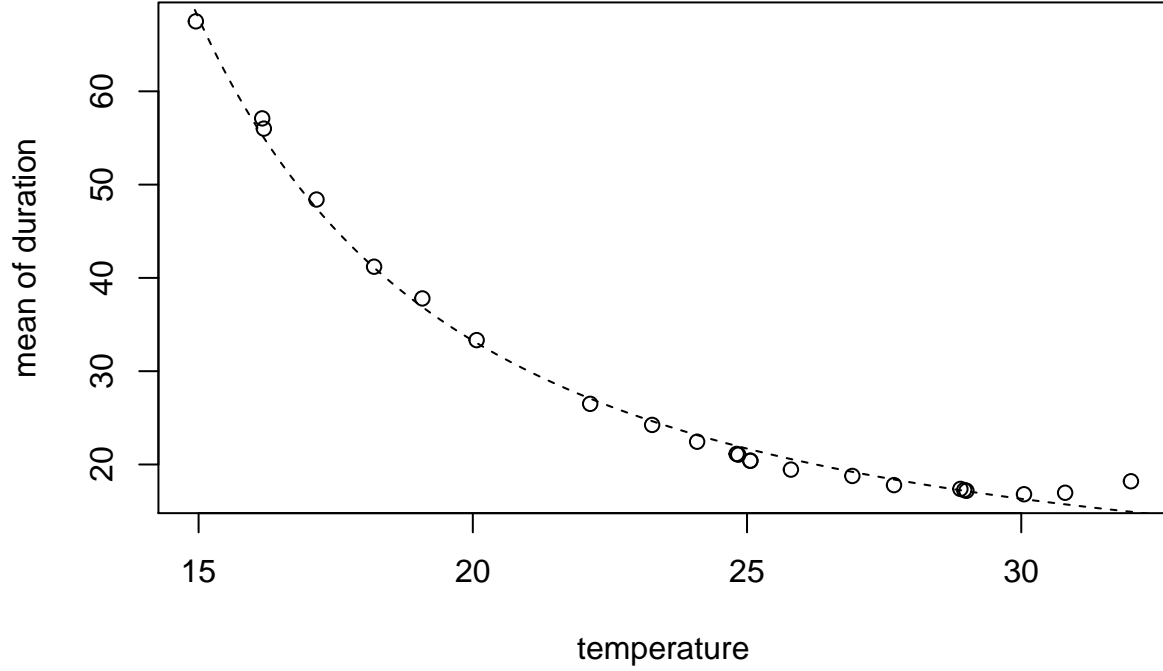

Answer:

There is a quite obvious trend in the Pearson residuals plot, where the Pearson residuals first decreases and then increases. The correlated and not normally distributed residuals indicates some problems existing in the model structure, thus it is not a good fit model in terms of the distribution of residuals.

(iv)

```r
# plot the curve of fitted values on the plot of the data
plot(duration~temp, fly, xlab="temperature", ylab="mean of duration")
x = seq(10,40,0.001)
lines(x, predict(mod1, data.frame(temp=x), type="response"), lty=2)
```

Answer:

Most of the data points with low temperatures lie in the estimated curve with a good fit. However, for large values of temperature, especially for those over 30, the data points does not fit in the estimated curve. This also indicates that some changes are needed for the current model.

(v)

Before the two changes, our model is: $\log \mu_i = \beta_0 + \beta_1 \frac{1}{temp_i}$, where $\mu_i = E(y_i)$

Now we first add an additional predictor "temp", so our model becomes: $\log \mu_i = \beta_0 + \beta_1 \frac{1}{temp_i} + \beta_2 temp_i$

In the plot of "log(duration) ~ temp" in question (ii), the points first decreases then increases. Since that in our data, the values of "temp" are all positive, so the signs of β_1 and β_2 should be the same, i.e. $\beta_1 \beta_2 > 0$.

Therefore:

(1) when $\beta_1 > 0$, $\beta_2 > 0$, we have: $\log \mu_i = \beta_0 + \beta_1 \frac{1}{temp_i} + \beta_2 temp_i \geq \beta_0 + 2\sqrt{\beta_1 \beta_2}$, where the equality is reached when $\beta_1 \frac{1}{temp_i} = \beta_2 temp_i$, i.e. $temp_i = \sqrt{\beta_1/\beta_2}$

(2) when $\beta_1 < 0$, $\beta_2 < 0$, we have: $\log \mu_i = \beta_0 + \beta_1 \frac{1}{temp_i} + \beta_2 temp_i \leq \beta_0 - 2\sqrt{\beta_1 \beta_2}$, where the equality is also reached when $temp_i = \sqrt{\beta_1/\beta_2}$

In either cases, they are adding restrictions on the coefficients, which will reduce the degrees of freedoms and may not let the algorithm reach the optimal model. Therefore, by changing "1/temp" to "1/(temp-d)" we can eliminate these restrictions and also are able to find the best model by choosing an optimal value of constant d.

(vi)

```r
# find the optimal d with smallest deviance
d = seq(45,100,0.01)
dev = NULL
for (i in 1:length(d)) {
    model = glm(duration ~ I(1/(temp-d[i])) + temp,  fly,
                family=Gamma(link=log), weights=batch)
    dev[i] = deviance(model)
}

dev[which.min(dev)]
```

```
## [1] 0.3182334
```

```r
d_opt = d[which.min(dev)]; d_opt
```

```
## [1] 58.64
```

Answer:

At optimal $d^* = 58.64$, the Gamma GLM model gets the smallest devaince 0.3182334.

Thus, with the optiaml d^*, now our model is: $\log \mu_i = \beta_0 + \beta_1 \frac{1}{temp_i - 58.64} + \beta_2 temp_i$

```r
# fit the model with the optimal d
mod2 = glm(duration ~ I(1/(temp-d_opt)) + temp,  fly,
           family=Gamma(link=log), weights=batch)
summary(mod2)
```

```
##
## Call:
## glm(formula = duration ~ I(1/(temp - d_opt)) + temp, family = Gamma(link = log),
##     data = fly, weights = batch)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q        Max
## -0.23768  -0.09367  -0.01390   0.07567   0.21227
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.20172    0.04393   72.88   <2e-16 ***
## I(1/(temp - d_opt)) -217.00306    4.39198  -49.41   <2e-16 ***
## temp                  -0.26481    0.00366  -72.36   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.01591676)
##
##     Null deviance: 617.86435  on 22  degrees of freedom
## Residual deviance:   0.31823  on 20  degrees of freedom
## AIC: 418.98
##
## Number of Fisher Scoring iterations: 3
```

Answer:

The residual degrees of freedom is $n - p - 1 = 23 - 3 - 1 = 19$, where n is the number of observations, p is the number of parameters in this model, and 1 is the degree of freedom used by d^*.

To estimate the coefficient of variation σ of the Gamma distribution, we have: $\sigma^2 = 1/\nu$, and the dispersion parameter $\phi = 1/\nu$, so:

$$\hat{\sigma}^2 = \hat{\phi} = \frac{1}{\hat{\nu}} = \frac{X^2}{df}$$

where X^2 is the Pearson's chi-square statistic and df is the residual degrees of freedom, which is 19, not the same as shown in the model summary.

So we cannot use the dispersion parameter from the model summary directly.

```
# estimate sigma
sigma = sqrt(deviance(mod2,type="pearson") / 19);  sigma
```
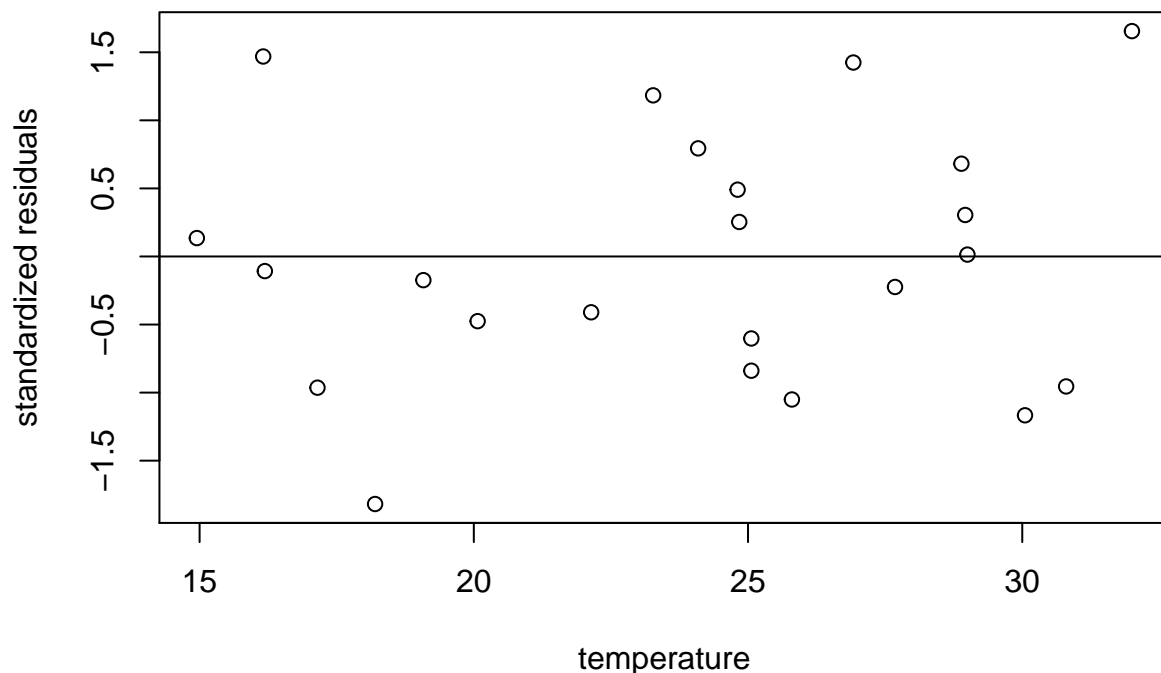
```
## [1] 0.1294184
```

Answer:

The estimated $\hat{\sigma}$ for this model is 0.1294184.

Thus, the standardized Pearson residuals $r_i = \frac{r_P}{\hat{\sigma}}$, where r_P is the Pearson residuals.

```
# plot standardized residuals ~ temp
resids = residuals(mod2, type="pearson") / sigma
plot(resids~temp, fly, xlab="temperature", ylab="standardized residuals")
abline(h=0)
```


Answer:

The residuals plot does not have any obvious trend any more, and it looks good and normally distributed with constant variance, which indicates that $mod2$ fits the data much better than $mod1$.

```
# plot the curve of fitted values on the plot of the data
plot(duration~temp, fly, xlab="temperature", ylab="mean of duration")
x = seq(10,40,0.001)
lines(x, predict(mod2, data.frame(temp=x), type="response"), lty=2)
```
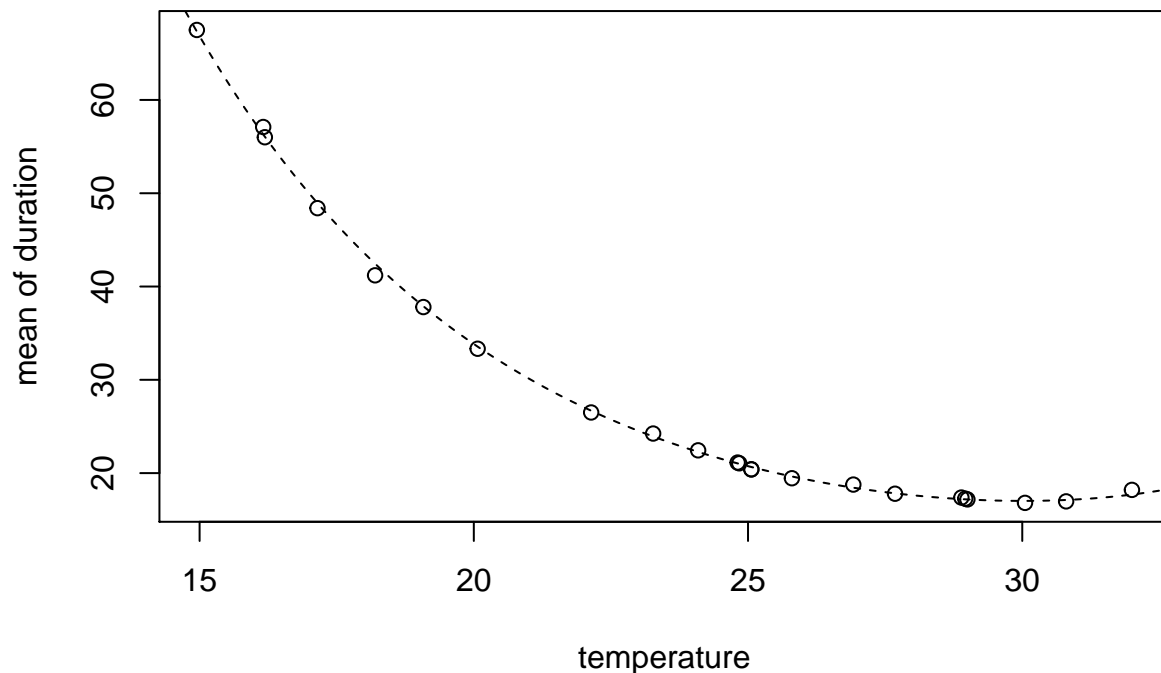

Answer:

Different from $mod1$, now we see that almost all the data points lie in the estimated curve with a good fit, both for the points in the range of low temperatures and high temperatures. Thus $mod2$ fits the data much better than $mod1$.

(vii)

The linear model with log transformed response here is:
linear model: $\log duration_i = \beta_0 + \beta_1 \frac{1}{temp_i - 58.64} + \beta_2 temp_i + \epsilon_i$

```
# fit a log transformed linear model
mod3 = lm(log(duration) ~ I(1/(temp-d_opt)) + temp,  fly)

# compare the coefficients of mod2 & mod3
summary(mod3)$coefficients
```

```
##                        Estimate  Std. Error   t value     Pr(>|t|)
## (Intercept)           3.1430273 0.045212638  69.51656 2.496957e-25
## I(1/(temp - d_opt)) -221.6556323 4.535174647 -48.87477 2.754473e-22
## temp                 -0.2680964 0.003794445 -70.65498 1.806626e-25
```

11

```r
summary(mod2)$coefficients
```

```
##                        Estimate  Std. Error   t value      Pr(>|t|)
## (Intercept)           3.2017231 0.043929915  72.88252 9.733099e-26
## I(1/(temp - d_opt)) -217.0030623 4.391979479 -49.40894 2.220050e-22
## temp                 -0.2648064 0.003659673 -72.35794 1.124007e-25
```

Answer:

We see that both the estimated coefficients and the standard errors of the coefficients are very similar to each other in the Gamma GLM model and the log transformed linear model. This is consistent with the conclusions in question (e).