# Task Solution Chapter 3

*Witold Wolski*

*December 3, 2016*

## Task 1

verify formula for $var(Y)$

$$var(Y) = \frac{\phi}{w} b''(\theta)$$

$$l(\theta, \phi, y) = \frac{y\theta - b(\theta)}{\phi/w} + c(y, \phi)$$

and using proofs in 3.3.2.

To warm up lets show that the mean is equal to $b'(\theta)$.

$$\begin{aligned}
0 &= E(\frac{\partial l}{\partial \theta}) \\
&= E(\frac{y - b'(\theta)}{\phi/w}) = w/\phi E(y - b'(\theta)) = w/\phi(E(y) - b'(\theta)) \\
&\rightarrow E(Y) = b'(\theta)
\end{aligned}$$

According to Result 2 on page 64

with:

$$\begin{aligned}
-E(\frac{\partial^2 l}{\partial \theta^2}) &= E(\frac{\partial l}{\partial \theta}^2) \\
-E(-wb''(\theta)/\phi) &= E((\frac{y - b'(\theta)}{\phi/w})^2) \\
w/\phi E(b''(\theta)) &= (w/\phi)^2 E((y - E(Y))^2) \\
\phi/w E(b''(\theta)) &= E((y - E(Y))^2) \\
\frac{\phi}{w} b''(\theta) &= Var(Y)
\end{aligned}$$

with $Var(Y) = E((y - E(Y))^2)$ and $E(b''(\theta)) = b''(\theta)$ (but this seems quite a stretch).

## Task 2

Use equations (3.8) and (3.9) to verify the entries in the table above.

Equation (3.8) is:

$$\mu = E(Y) = b'(\theta)$$

Equation (3.9) is:

$$var(Y) = \frac{\phi}{w} b''(\theta)$$

**i**

For normal distribution:

$$b(\theta) = \theta^2/2 \rightarrow b'(\theta) = \theta \rightarrow b''(\theta) = 1$$

**ii**

For Poisson distribution

$$b(\theta) = e^{\theta} \rightarrow b'(\theta) = e^{\theta} \rightarrow b''(\theta) = e^{\theta}$$

**iii**

For Binomial

$$b(\theta) = \log(1 + e^{\theta}) \rightarrow b'(\theta) = \frac{e^{\theta}}{1 + e^{\theta}} \rightarrow b''(\theta) = \frac{e^{\theta}}{e^{\theta} + 1} - \frac{e^{2\theta}}{(e^{\theta} + 1)^2}$$

Need to check how to differentiate this.

**iv**

For Gamma

$$b(\theta) = -log(-\theta) \rightarrow b'(\theta) = +1/\theta \rightarrow b''(\theta) = 1/\theta^2 = (-1\theta)^2 = \mu^2$$

## Task 3

Task 3 Verify the canonical links for the Normal, Binomial, Poisson and Gamma.
The link function $g$ for which $\theta_i = x_i^T \beta$ (that is, for which $g = (b')^{-1}$) is called the canonical link.
Hence, the task is to find the inverse of $g = (b')^{-1}$.

**for Normal**

$$b'(\theta) = \theta \rightarrow \theta = g(\mu) = 1(\mu)$$

**for Poisson**

$$b'(\theta) = e^{\theta} \rightarrow \mu = e^{\theta} \rightarrow \theta = \log(\mu)$$

**for Binomial**

$$b'(\theta) = \log(1 + e^{\theta}) \rightarrow \mu = \log(1 + e^{\theta}) \rightarrow \theta = \log(\mu_i/(1 - \mu_i))$$

**For Gamma**

$$b'(\theta) = -1/\theta \rightarrow \mu = -1/\theta \rightarrow \theta = g(\mu) = -1/\mu$$

## Task 4

For the canonical link (see, in particular, equation (3.10)) show that

$$
\begin{aligned}
\frac{\partial^2 l}{\partial\beta_j\partial\beta_k} &= \\
&= \frac{1}{\phi}\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}(\sum_i\{w_i[y_i\theta_i - b(\theta_i)]\}) \\
&= \frac{1}{\phi}\sum_i\{-x_{ij}w_i h'(x_i^T\beta)x_{ik}\}
\end{aligned}
$$

Hint :

$$\sum w_i y_i \frac{\partial\theta}{\partial\beta} = \sum w_i \frac{\partial b(\theta)}{\partial\beta}$$

$$b'(\theta) = h(x^T\beta) \to b(\theta) = \int h(x^T\beta) + C$$

than:

$$
\begin{aligned}
&= \frac{1}{\phi}\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}(\sum_i\{w_i[y_i\theta_i - b(\theta_i)]\}) \\
&= \frac{1}{\phi}\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}(\sum_i\{w_i[y_i\theta_i - (\int h(x^T\beta) + C)]\})
\end{aligned}
$$

## Task 5

Show that, for the maximal model, $y_i$ is the maximum likelihood estimate of $\mu_i$. (Hint: the likelihood equation for $\theta_i$ is $y_i = b'(\theta_i)$, and we already know that $b'(\theta_i) = \mu_i$ .)

$$y_i = b'(\theta_i) = \mu_i$$

$$l(\mu; y, \phi) = \sum_i \frac{y_i\theta - b(\theta)}{\phi/w_i} + c(y, \phi)$$

For the maximal model :

$$
\begin{aligned}
&\partial l/\partial\theta_i = (w_i/\phi)\{y_i - b'(\theta_i)\} = 0 \\
&\to y_i = b'(\theta_i) = \mu_i
\end{aligned}
$$

for other models:

$$\partial l/\partial\theta_i = \sum_i (w_i/\phi)\{y_i - b'(\theta_i)\} = 0$$

The model that allows the mean of each observation to be a separate parameter is called the **full**, **saturated** or **maximal** model.

The maximum possible value of $L$ as the components of $\mu$ vary without any restrictions (or, essentially equivalently, when there is an element of $\beta$ for each observation, so that $\mu_i^\diamond = h(\eta i) = h(\beta i)$ occurs when $\mu_i^\diamond = y_i$, hence $\mu_i^\diamond = y_i$ for all $i$.

# Task 6

Deviances for common distributions.

Confirm results using the table in 3.3.4 and equation (3.12).

Eq 3.12 is:

$$D(y, \hat{\mu}) = \phi S(y, \hat{\mu}) = 2 \sum_i w_i [y_i(\hat{\theta}_i^\diamond - \hat{\theta}_i) - b(\hat{\theta}_i^\diamond) + b(\hat{\theta}_i)]$$

### For Normal

$\hat{\theta^\diamond} = y_i$ and $\hat{\theta} = \mu$; $b(\theta) = \theta^2/2$ than $b(\hat{\theta^\diamond}) = y_i^2/2$ and $b(\hat{\theta}) = \mu^2/2$; and $w_i = 1$.

$$\begin{aligned}
2 \sum_i [y_i(y_i - \mu) - y_i^2/2 + \mu^2/2] \\
&= 2 \sum_i [y_i^2 - y_i\mu - y_i^2/2 + \mu^2/2] \\
&= \sum_i [y_i^2 - 2y_i\mu + \mu^2] \\
&= \sum_i (y_i - \mu)^2
\end{aligned}$$

### For Poisson

$\mu = e^\theta \to \theta = \log(\mu)$. Than

$\hat{\theta} = \log(\mu)$ and $\hat{\theta^\diamond} = \log y_i$; $b(\theta) = e^\theta$ than

$b(\hat{\theta^\diamond}) = e^{\log(y_i)} = y_i$ and $b(\hat{\theta}) = e^{\log(\mu)} = \mu$ and $\phi = 1$, $w_i = 1$

$$2 \sum_i \{y_i(\log(y_i) - \log(\mu)) - y_i + \mu\} = 2 \sum_i \{y_i \log(y_i/\mu) - y_i + \mu\}$$

## For Binomial

TODO

# Task 7

Use the results in 3.6.4 to obtain simple expressions for the null deviance for Poisson (weights must be one) and Gamma (assume weights are one).

If $W = \sum w_i$

$$\mu = \sum w_i y_i / W$$

For Poisson and Gamma model $w_i = 1$, than $\mu = 1/N \sum y_i$

**For Poison**

$$\begin{aligned}
2 \sum y_i \log(y_i/\hat{\mu}_i) - (y_i - \hat{\mu}_i) = \\
&= 2 \sum (y_i \log(\frac{y_i}{1/N \sum y_i}) - (y_i - 1/N \sum y_i)) \\
&= 2 \sum (y_i \log(y_i) - \log(1/N \sum y_i) - (y_i) + (1/N \sum y_i)) \\
&= 2 \sum y_i (\log(y_i) - \log(1/N \sum y_i))
\end{aligned}$$

**For Gamma**

Similarly, TODO

# Task 8

For the Poisson distribution, verify that $D$ and $\chi^2$ are asymptotically equivalent. (Hint: use that for small $|x - y|, x \log(x/y) \approx (x - y) + (x - y)^2/(2y)$.)

$$\chi^2 = \sum e_{P,i}^2 = \sum (\sqrt{w_i} \frac{y_i - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)}})^2$$

for Poisson $w_i = 1$ and $Var(Y_i) = \frac{\phi}{w_i} V(\mu_i) = \frac{\phi}{w_i} b''(\theta)$ and $b''(\theta) = V(\mu_i) = \mu$,

than:

$$\chi^2 = \sum e_{P,i}^2 = \sum (\frac{y_i - \hat{\mu}_i}{\sqrt{\mu_i}})^2$$

$$\begin{aligned}
D(y_i, \mu_i) = 2 \sum y_i \log(y_i/\hat{\mu}_i) - (y_i - \hat{\mu}_i) = \\
&= 2 \sum (y_i - \hat{\mu}_i) + (y_i - \hat{\mu}_i)^2/(2\hat{\mu}_i) - (y_i - \hat{\mu}_i) \quad = \\
&= \sum ((y_i - \hat{\mu}_i)^2/\hat{\mu}_i)
\end{aligned}$$

# Task 9

Obtain the form of the **Pearson residual** for the Binomial, Poisson and Gamma. (These are given in Chapters 4 and 5)

$$e_{P,i} = w_i \frac{y_i - \hat{y}_i}{\sqrt{V(\hat{\mu_i})}}$$

**For Binomial**

$$Var(\mu_i) = \phi/w_i b''(\theta) = (1/n)\mu(1-\mu)$$

than

$$e_{p,i} = \frac{y_i - \mu_i}{Var(\mu_i)} \overset{with1}{=} \frac{y_i - \mu_i}{\sqrt{\mu_i(1-\mu_i)/n_i}}$$

**For Poisson**

$$Var(\mu_i) = \phi/w_i b''(\theta) = (1/1)\mu$$

than

$$e_{p,i} = \frac{y_i - \mu_i}{\sqrt{\mu_i}}$$

**For Gamma**

$$Var(\mu_i) = \phi/w_i b''(\theta) = (\phi/1)\mu^2$$

than

$$e_{p,i} = \frac{y_i - \mu_i}{\sqrt{\phi}\mu_i}$$

# Task 10

- How does a GLM with identity link relate to a linear model?

It's the same as a linear model.

- Under which situations does it make sense to fit a linear model instead of a GLM with identity link?

Almost always when the response $y$ is continuous and normally distributed.

# Task 11

Confirm that:

$$\chi^2 = \sum e_{P,i}^2 = \sum \{ \frac{[s_i - n_i\hat{\mu_i}]^2}{n_i\hat{\mu_i}} + \frac{[(n_i - s_i) - n_i(1-\hat{\mu_i})]^2}{n_i(1-\hat{\mu_i})} \}$$

with

$$e_{P,i} = \frac{y_i - \hat{\mu_i}}{\sqrt{\hat{\mu_i}(1-\hat{\mu_i})/n_i}}$$

and $s_i = n_i y_i$.

use:

$$\frac{1}{p(1-p)} = \frac{1}{p} + \frac{1}{1-p}$$