P8131 Spring 2022 Homework #1 Solution

1. Show that the following distributions belong to the exponential family. Find the natural parameter θ, scale parameter ϕ and convex function $b(\theta)$. Also find the EY and $\mathrm{Var}(Y)$ as functions of the natural parameter. Specify the canonical link functions.

   (a) Exponential distribution $Exp(\lambda)$, $f(y;\lambda) = \lambda e^{-\lambda y}$;

   (b) Binomial distribution $Bin(n,\pi)$, $f(y;\pi) = \binom{n}{y}\pi^y(1-\pi)^{n-y}$, where n is known;

   (c) Poisson distribution $Pois(\lambda)$, $f(y;\lambda) = \frac{1}{y!}\lambda^y e^{-\lambda}$;

   (d) Negative binomial distribution $NB(m,\beta)$, $f(y;\beta) = \binom{y+m-1}{m-1}\beta^m(1-\beta)^y$, where m is known;

   (e) The Gamma distribution $Gamma(\alpha,\beta)$, $f(y;\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\beta y}$, where the shape parameter α is known.

   Solution:

   (a) $\theta = -\lambda, \phi = 1, b(\theta) = -\log(-\theta), E(Y) = -\frac{1}{\theta}, Var(Y) = \frac{1}{\theta^2}$.
   Canonical link: $g(\mu) = -\frac{1}{\mu}$.

   (b) $\theta = log\frac{\pi}{1-\pi}, \phi = 1, b(\theta) = n\log(1+e^\theta), E(Y) = \frac{ne^\theta}{1+e^\theta}, Var(Y) = \frac{ne^\theta}{(1+e^\theta)^2}$.
   Canonical link: $g(\mu) = \log(\frac{\mu}{n-\mu})$.

   (c) $\theta = \log\lambda, \phi = 1, b(\theta) = e^\theta, E(Y) = e^\theta, Var(Y) = e^\theta$.
   Canonical link: Log link $g(\mu) = \log(\mu)$

   (d) $\theta = \log(1-\beta), \phi = 1, b(\theta) = -m\log(1-e^\theta), E(Y) = \frac{me^\theta}{1-e^\theta}, Var(Y) = \frac{me^\theta}{(1-e^\theta)^2}$.
   Canonical link: $g(\mu) = \log(\frac{\mu}{m+\mu})$

   (e) $\theta = -\beta, \phi = 1, b(\theta) = -\alpha\log(-\theta), E(Y) = -\frac{\alpha}{\theta}, Var(Y) = \frac{\alpha}{\theta^2}$.
   Canonical link: $g(\mu) = -\frac{\alpha}{\mu}$.

2. Assume $Y_1, Y_2, ..., Y_n$ are independent and follow a binomial distribution where $Y_i \sim Bin(m,\pi_i)$ and m is known. Furthermore, assume $\log\frac{\pi_i}{1-\pi_i} = X_i\beta$. What are the expressions of deviance residuals and Pearson residuals respectively (use $\hat\beta$ to represent the MLE)? What are the expressions of the deviance and Pearson's χ^2 statistic?
   Solution:
   Deviance residual:

$$r_{D_i} = sign(y_i - m\hat\pi_i)\sqrt{D(y_i, m\hat\pi_i)} = sign(y_i - m\hat\pi_i)\sqrt{2[l(y_i,y_i) - l(y_i, m\hat\pi_i)]}$$

$$= sign(y_i - m\hat\pi_i)\sqrt{2\left[y_i\log\frac{y_i}{m\hat\pi_i} + (m-y_i)\log\frac{m-y_i}{m(1-\hat\pi_i)}\right]}.$$

Deviance statistic:

$$D = \sum_{i=1}^{n} r_{D_i}^2 = 2\sum_{i=1}^{n}\left[y_i \log\frac{y_i}{m\hat{\pi}_i} + (m - y_i)\log\frac{m - y_i}{m(1-\hat{\pi}_i)}\right].$$

Pearson residual:

$$r_{P_i} = \frac{y_i - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)}} = \frac{y_i - m\hat{\pi}_i}{\sqrt{m\hat{\pi}_i(1-\hat{\pi}_i)}}.$$

Pearson's χ^2 statistic:

$$G = \sum_{i=1}^{n}\frac{(y_i - m\hat{\pi}_i)^2}{m\hat{\pi}_i(1-\hat{\pi}_i)},$$

where

$$\hat{\pi}_i = \frac{\exp(x_i\hat{\beta})}{1+\exp(x_i\hat{\beta})}$$

3. Consider the binary response variable $Y \sim$ *Bernoulli* with $P(Y=1) = \pi$ and $P(Y=0) = 1-\pi$. Observations Y_i, $i = 1, \ldots, n$, are independent and identically distributed as Y.

   (a) Find the Wald test statistic, the score test statistic, and the likelihood ratio test statistic to test hypothesis $H_0 : \pi = \pi_0$.

   (b) With large samples, the Wald test statistic, score test statistic and the likelihood ratio test statistic approximately have the $\chi^2(1)$ distribution. For $n = 10$ and data (0, 1, 0, 0, 1, 0, 0, 0, 1, 0), use these statistics to test null hypotheses on for (i) $\pi_0 = 0.1$, (ii) $\pi_0 = 0.3$, (iii) $\pi_0 = 0.5$.

   (c) Do the Wald test, score test, and the likelihood ratio test lead to the same conclusions in (b)?

Solution:

(a) To test $H_0 : \pi = \pi_0$ and $H_1 : \pi \neq \pi_0$, calculate

$$\begin{aligned}
l(\pi) &= \sum_{i=1} y_i \log\pi + \sum_{i=1}(1-y_i)\log(1-\pi) \\
l'(\pi) &= \frac{\sum_{i=1}^n y_i}{\pi} - \frac{n - \sum_{i=1}^n y_i}{1-\pi} \\
l''(\pi) &= -\frac{\sum_{i=1}^n y_i}{\pi^2} - \frac{n - \sum_{i=1}^n y_i}{(1-\pi)^2} \\
\mathcal{I}(\pi) &= E(-l''(\pi)) = \frac{n}{\pi(1-\pi)}.
\end{aligned}$$

Set $l'(\pi) = 0$ to solve for the MLE $\hat{\pi}$ and check for the second derivative, the MLE $\hat{\pi} = \sum_{i=1}^n y_i/n$.
Wald test statistics:

$$TS_W = (\hat{\pi} - \pi_0)^T \mathcal{I}(\hat{\pi})(\hat{\pi} - \pi_0) = \frac{(\hat{\pi} - \pi_0)^2 n}{\hat{\pi}(1-\hat{\pi})}.$$

Score test statistics:

$$
\begin{aligned}
TS_S &= s(\pi_0)^T \mathcal{I}^{-1}(\pi_0) s(\pi_0) \\
&= \frac{\frac{\sum_{i=1}^n y_i - n\pi_0}{\pi_0(1-\pi_0)}}{\frac{n}{\pi_0(1-\pi_0)}} \\
&= \frac{(\sum_{i=1}^n y_i - n\pi_0)^2}{n\pi_0(1-\pi_0)} \\
&= \frac{(\hat{\pi} - \pi_0)^2 n}{\pi_0(1-\pi_0)}.
\end{aligned}
$$

Likelihood ratio test statistics:

$$
\begin{aligned}
TS_{LR} &= 2[l(y,\hat{\pi}) - l(y,\pi_0)] \\
&= 2[\sum_{i=1} y_i \log \hat{\pi} + \sum_{i=1} (1-y_i)\log(1-\hat{\pi}) - \sum_{i=1} y_i \log \pi_0 - \sum_{i=1}(1-y_i)\log(1-\pi_0))] \\
&= 2[\sum_{i=1} y_i \log(\frac{\hat{\pi}}{\pi_0}) + (n - \sum_{i=1} y_i)\log(\frac{1-\hat{\pi}}{1-\pi_0})].
\end{aligned}
$$

(b) $\hat{\pi} = \sum_{i=1} y_i/n = 0.3$. $\alpha = 0.05$. The critical value $\chi^2(1)_{0.05}$ is 3.84.

i. $\pi_0 = 0.1$
Wald test: $TS_W = 1.9 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.
Score test: $TS_S = 4.44 > \chi^2(1)_{0.05} = 3.84$, reject H_0 at 5% level of significance.
Likelihood ratio test: $TS_{LR} = 3.07 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.

ii. $\pi_0 = 0.3$
Wald test: $TS_W = 0 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.
Score test: $TS_S = 0 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.
Likelihood ratio test: $TS_{LR} = 0 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.

iii. $\pi_0 = 0.5$
Wald test: $TS_W = 1.9 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.
Score test: $TS_S = 1.6 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.
Likelihood ratio test: $TS_{LR} = 1.65 < \chi^2(1)_{0.05} = 3.84$, fail to reject H_0 at 5% level of significance.

(c) No, they do not necessarily lead to the same conclusion. For example, when testing $\pi_0 = 0.1$, Score test rejects the null hypothesis, but Wald test and Likelihood ratio test fail to reject the null hypothesis at 5% level of significance.

4. Assume $Y_i \sim Pois(\lambda), i = 1, ..., n$. We are interested in testing $H_0 : \log\lambda = \log\lambda_0$. What are the Wald test statistic, the score test statistic, and the likelihood ratio test statistic?

Solution:

Let $\log\lambda = \theta$ and $\log\lambda_0 = \theta_0$.

$$f_\lambda(y_i) = e^{-\lambda}\lambda^{y_i}/y_i!$$

$$l(\lambda) = -n\lambda + \log(\lambda)\sum_{i=1}^n y_i - \sum_{i=1}^n \log(y_i!)$$

$$l'(\lambda) = -n + \frac{1}{\lambda}\sum_{i=1}^n y_i$$

$$l''(\lambda) = -\frac{1}{\lambda^2}\sum_{i=1}^n y_i$$

$$\mathcal{I}(\lambda) = E(-l''(\lambda)) = \frac{n}{\lambda} = E\left(\frac{dl(\lambda)}{d\lambda}\right)^2$$

$$s(\theta) = \frac{dl(\lambda)}{d\lambda}\frac{d\lambda}{d\theta} = (-n + \frac{1}{e^\theta}\sum_{i=1}^n y_i)e^\theta = \sum_{i=1}^n y_i - ne^\theta$$

$$\mathcal{I}(\theta) = E\left(\frac{dl(\lambda)}{d\lambda}\frac{d\lambda}{d\theta}\right)^2 = \mathcal{I}(\lambda)e^{2\theta} = \frac{ne^{2\theta}}{e^\theta} = ne^\theta$$

Note: You can also rewrite the likelihood function in terms of theta, and then directly calculate the score function and fisher information as functions of theta.

By invariance property of MLE, we know $\hat{\theta} = \log(\hat{\lambda}) = \log(\bar{Y})$.

For testing $H_0 : \log\lambda = \log\lambda_0$, i.e., $H_0 : \theta = \theta_0$,

Wald test statistic:

$$\begin{aligned} TS_W &= (\hat{\theta} - \theta_0)^2\mathcal{I}(\hat{\theta}) \\ &= \Big(\log(\bar{y}) - \log(\lambda_0)\Big)^2 n\bar{y} \end{aligned}$$

Score test statistics:

$$\begin{aligned} TS_S &= s(\theta_0)^2/\mathcal{I}(\theta_0) \\ &= \frac{(\sum_{i=1}^n y_i - n\lambda_0)^2}{n\lambda_0} = \frac{n(\bar{y} - \lambda_0)^2}{\lambda_0} \end{aligned}$$

Likelihood ratio test statistics:

$$\begin{aligned} TS_{LR} &= 2[l(y, \hat{\theta}) - l(y, \theta_0)] \\ &= 2[n(e_0^\theta - e^{\hat{\theta}}) + \sum_{i=1}^n y_i(\log(e^{\hat{\theta}}) - \log(e_0^\theta))] \\ &= 2[n(\lambda_0 - \bar{y}) + \sum_{i=1}^n y_i \log\frac{\bar{y}}{\lambda_0}] \end{aligned}$$

Note: $\hat{\lambda} = \bar{Y}$ is the MLE of λ.