UNIVERSITY OF EDINBURGH
SCHOOL OF MATHEMATICS

Generalised Regression Models

GRM: Solutions 1 **Semester 1, 2022–2023**

---

1. The least squares estimates are obtained by minimising

$$Q = \sum_{i=1}^{n} (y_i - \mathrm{E}(Y_i|x_i))^2 = (y_1 - \theta)^2 + (y_2 - 2\theta + \phi)^2 + (y_3 - \theta - 2\phi)^2$$

with respect to θ and ϕ.

Differentiating Q with respect to θ and ϕ gives

$$\frac{\partial Q}{\partial \theta} = -2(y_1 - \theta) - 4(y_2 - 2\theta + \phi) - 2(y_3 - \theta - 2\phi) = 2(-y_1 + \theta - 2y_2 + 4\theta - 2\phi - y_3 + \theta + 2\phi)$$

$$= 2(6\theta - y_1 - 2y_2 - y_3)$$

$$\frac{\partial Q}{\partial \phi} = 2(y_2 - 2\theta + \phi) - 4(y_3 - \theta - 2\phi) = 2(5\phi + y_2 - 2y_3)$$

Solving $\frac{\partial Q}{\partial \theta} = \frac{\partial Q}{\partial \phi} = 0$ yields the least squares estimates of θ and ϕ:

$$\widehat{\theta} = \frac{1}{6}(y_1 + 2y_2 + y_3) \quad \text{and} \quad \widehat{\phi} = \frac{1}{5}(2y_3 - y_2).$$

2. The least squares estimates are obtained by minimizing the function

$$Q = \sum_{i=1}^{n} (y_i - \mathrm{E}(Y_i|x_i))^2 = \sum_{i=1}^{n} (y_i - \gamma - \beta(x_i - \bar{x}))^2$$

Differentiating Q with respect to γ and β, and equating each of these equations to zero gives the *normal equations* for determining the least squares estimates $\widehat{\gamma}$ and $\widehat{\beta}$:

$$\begin{aligned} \frac{\partial Q}{\partial \gamma} &= -2\sum_{i=1}^{n} (y_i - \gamma - \beta(x_i - \bar{x})) = 0 \\ \frac{\partial Q}{\partial \beta} &= -2\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \gamma - \beta(x_i - \bar{x})) = 0 \end{aligned}$$

Solving for $\widehat{\gamma}$ and $\widehat{\beta}$ gives:

$$\widehat{\gamma} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}, \qquad \widehat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}},$$

where $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \bar{x})x_i = \sum_{i=1}^{n} x_i^2 - n^{-1}\left(\sum_{i=1}^{n} x_i\right)^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$,
and $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum(x_i - \bar{x})y_i = \sum x_i y_i - n\bar{x}\bar{y}$.

Variances: $\mathrm{var}(\widehat{\gamma}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{var}(Y_i) = \frac{\sigma^2}{n}$, $\mathrm{var}(\widehat{\beta}) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2\mathrm{var}(Y_i)}{S_{xx}^2} = \frac{S_{xx}\sigma^2}{S_{xx}^2} = \frac{\sigma^2}{S_{xx}}$

Covariance: $\mathrm{cov}(\widehat{\gamma}, \widehat{\beta}) = \mathrm{cov}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i, \frac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{S_{xx}}\right) = \frac{1}{nS_{xx}}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_j - \bar{x})\mathrm{cov}(Y_i, Y_j) = 0$
as $\mathrm{cov}(Y_i, Y_j) = 0$ $(i \neq j)$, and $\mathrm{cov}(Y_i, Y_i) = \mathrm{var}(Y_i)$ thus $\sum_{i=1}^{n}(x_i - \bar{x})\mathrm{var}(Y_i) = \sigma^2\sum_{i=1}^{n}(x_i - \bar{x}) = 0$.

For the data set: $n = 18$, $\bar{x} = 20$, $\bar{y} = 5$, $S_{xx} = \sum x^2 - n\bar{x}^2 = 18(3456) - 18(20)^2 = 55008$,
$S_{yy} = \sum y^2 - n\bar{y}^2 = 18(352) - 18(5)^2 = 5886$, $S_{xy} = \sum xy - n\bar{x}\bar{y} = 18(576) - 18(20)(5) = 8568$

(a) Test $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$.

$$\widehat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{8568}{55008} = 0.15576$$

$$\widehat{\sigma}^2 = \frac{1}{n-2}(S_{yy} - \frac{S_{xy}^2}{S_{xx}}) = \frac{1}{16}(5886 - \frac{8568^2}{55008}) = 284.47$$

$$\text{ESE}(\widehat{\beta}) = \sqrt{\text{var}(\widehat{\beta})} = \sqrt{\frac{\widehat{\sigma}^2}{S_{xx}}} = \sqrt{\frac{284.47}{55008}} = 0.07191$$

The test statistics is

$$t = \frac{\widehat{\beta} - 0}{ESE(\widehat{\beta})} = \frac{0.15576}{0.07191} = 2.166.$$

This is compared with $t_{16}(2.5\%) = 2.120$ for a two-sided 5% test. Thus, we can reject the null hypothesis H_0 at the 5% level, and conclude that $\beta \neq 0$. Note that the test is just significant at the 5% significance level [so the p-value (*significance probability*) of the test is about 5%].

(b) A 95% confidence (prediction) interval for a future observation at $x = x^* = 34$ is

$$\widehat{\gamma} + \widehat{\beta}(x^* - \bar{x}) \pm t_{n-2}(2.5\%)\sqrt{\widehat{\sigma}^2\left(1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}\right)}$$

i.e.

$$5 + 0.15576(34-20) \pm 2.120\sqrt{284.47\left(1 + \frac{1}{18} + \frac{(34-20)^2}{55008}\right)} = 7.181 \pm 36.798 = (-29.6, 44.0)$$

3. $RSS = \sum_{i=1}^n (Y_i - \widehat{Y}_i)^2 = \sum_{i=1}^n (Y_i - \widehat{\gamma} - \widehat{\beta}(x_i - \bar{x}))^2 = \sum_{i=1}^n Y_i^2 + n\widehat{\gamma}^2 + \widehat{\beta}^2 \sum_{i=1}^n (x_i - \bar{x})^2 - 2\widehat{\gamma}\sum_{i=1}^n Y_i - 2\widehat{\beta}\sum_{i=1}^n Y_i(x_i - \bar{x}) = \sum_{i=1}^n Y_i^2 - n\widehat{\gamma}^2 - \widehat{\beta}^2 \sum_{i=1}^n (x_i - \bar{x})^2$ (using $\widehat{\gamma} = \bar{y}$, $\widehat{\beta} = \frac{S_{xy}}{S_{xx}}$ and $\sum(x_i - \bar{x}) = 0$).

Therefore, $\text{E}(RSS) = \sum_{i=1}^n \text{E}(Y_i^2) - n\text{E}(\widehat{\gamma}^2) - \text{E}(\widehat{\beta}^2)\sum_{i=1}^n (x_i - \bar{x})^2$.

But

- $\sum \text{E}(Y_i^2) = \sum(\text{var}(Y_i) + \{\text{E}(Y_i)\}^2) = \sum(\sigma^2 + (\gamma + \beta(x_i - \bar{x}))^2) = n\sigma^2 + n\gamma^2 + \beta^2\sum(x_i - \bar{x})^2$
- $n\text{E}(\widehat{\gamma}^2) = n(\text{var}(\widehat{\gamma}) + \{\text{E}(\widehat{\gamma})\}^2) = n(\frac{\sigma^2}{n} + \gamma^2)$ as $\text{E}(\widehat{\gamma}) = \gamma$
- $\text{E}(\widehat{\beta}^2) = \text{var}(\widehat{\beta}) + \{\text{E}(\widehat{\beta})\}^2 = \frac{\sigma^2}{S_{xx}} + \beta^2$ as $\text{E}(\widehat{\beta}) = \beta$

$$\text{E}(RSS) = \left[n\sigma^2 + n\gamma^2 + \beta^2\sum(x_i - \bar{x})^2\right] - \left[n(\frac{\sigma^2}{n} + \gamma^2)\right] - \left[\frac{\sigma^2}{\sum(x_i - \bar{x})^2} + \beta^2\right]\sum(x_i - \bar{x})^2 = (n-2)\sigma^2.$$

Therefore, $\widehat{\sigma}^2 = \frac{RSS}{n-2}$ is an unbiased estimator of σ^2.

4. (a) For the linear regression model in which responses Y_i are uncorrelated with expectations βx_i and common variance σ^2, the least squares estimate of β minimizes the sum of squares

$$Q = \sum_{i=1}^n (y_i - \beta x_i)^2 = \sum y^2 - 2\beta\sum xy + \beta^2\sum x^2,$$

Differentiation with respect to β gives

$$\frac{\partial Q}{\partial \beta} = 2(\beta\sum x^2 - \sum xy),$$

so the least squares estimate of β is given by $\widehat{\beta} = \frac{\sum xy}{\sum x^2}$, as required.

(b) Given the value of $\mathbf{x} = [x_1 \dots x_n]^T$, the corresponding estimator of β has expectation and variance

$$
\begin{aligned}
\mathrm{E}(\widehat{\beta}\,|\,\mathbf{x}) &= \frac{\sum_i x_i\, \mathrm{E}(Y_i\,|\,x_i)}{\sum x^2} = \frac{\sum_i x_i\, \beta x_i}{\sum x^2} = \beta\,, \\
\mathrm{var}(\widehat{\beta}\,|\,\mathbf{x}) &= \frac{\sigma^2 \sum_i x_i^2}{(\sum x^2)^2} = \frac{\sigma^2}{\sum x^2}\,.
\end{aligned} \tag{1}
$$

(c) The alternative estimator $\widetilde{\beta} = \frac{\sum_i Y_i}{\sum_i x_i}$ has expectation and variance

$$
\begin{aligned}
\mathrm{E}(\widetilde{\beta}\,|\,\mathbf{x}) &= \frac{\sum_i \mathrm{E}(Y_i\,|\,x_i)}{\sum_i x_i} = \frac{\sum_i \beta x_i}{\sum_i x_i} = \beta\,, \\
\mathrm{var}(\widetilde{\beta}\,|\,\mathbf{x}) &= \frac{n\sigma^2}{(\sum_i x_i)^2} = \frac{\sigma^2}{(n\bar{x}^2)}\,.
\end{aligned}
$$

(d) From Hint,

$$
\sum_i x_i^2 - n^{-1}\left(\sum_i x_i\right)^2 = \sum_i (x_i - \bar{x})^2\,,
$$

which is non-negative, and hence $\sum x^2 \geq n\bar{x}^2$. Therefore the variance of $\widetilde{\beta}$ is at least as large as the variance of $\widehat{\beta}$ because

$$
\mathrm{var}(\widetilde{\beta}\,|\,\mathbf{x}) = \frac{\sigma^2}{n\bar{x}^2} \geq \frac{\sigma^2}{\sum x^2} = \mathrm{var}(\widehat{\beta}\,|\,\mathbf{x}).
$$

5. (a) If $\check{\beta}$ denotes $\sum_i a_i Y_i$, then

$$
\mathrm{E}(\check{\beta}\,|\,\mathbf{x}) = \sum_i a_i x_i \beta\,, \qquad \mathrm{var}(\check{\beta}\,|\,\mathbf{x}) = \sum_i a_i^2 \sigma^2\,.
$$

The expectation of $\check{\beta}$ equals β if and only if $\sum_i a_i x_i = 1$, i.e. iff $\mathbf{a}^T\mathbf{x} = 1$. To minimize the variance subject to the condition $\sum_i a_i x_i = 1$, we introduce a Lagrange multiplier λ, and find

$$
\frac{\partial}{\partial \mathbf{a}}\left(\mathbf{a}^T\mathbf{a} + \lambda\,\mathbf{a}^T\mathbf{x}\right) = 2\,\mathbf{a} + \lambda\,\mathbf{x}\,.
$$

Equating this to zero gives $\mathbf{a}$ as $-\frac{1}{2}\lambda\mathbf{x}$, and premultiplying $2\,\mathbf{a} + \lambda\,\mathbf{x} = \mathbf{0}$ by $\mathbf{x}^T$ gives

$$
\lambda = -2\,(\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{a} = -2\,(\mathbf{x}^T\mathbf{x})^{-1}.
$$

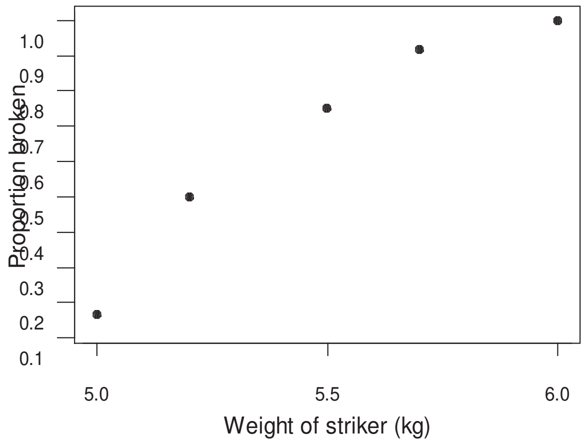
Hence, among unbiased estimators of the form $\sum_i a_i Y_i$, the variance is minimized when

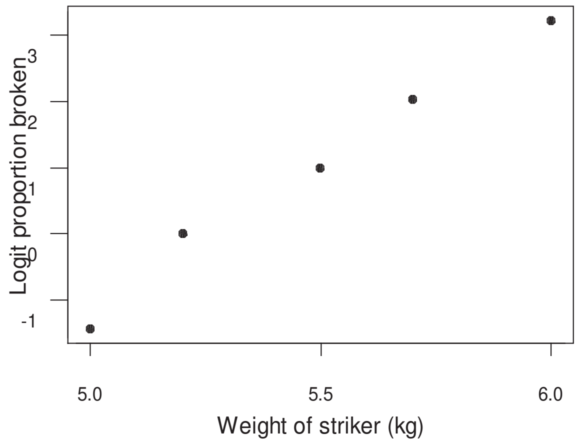
$$
\mathbf{a} = (\mathbf{x}^{\mathbf{T}}\mathbf{x})^{-\mathbf{1}}\mathbf{x}\,, \quad \text{or} \quad a_i = x_i / \sum x^2 \quad (i = 1, \dots, n)\,.
$$

(b) The least squares estimator, $\widehat{\beta} = \frac{\sum_i x_i Y_i}{\sum x^2}$, found in Question 4 may be written as $\widehat{\beta} = \sum_i a_i Y_i$, where $a_i = \frac{x_i}{\sum x^2}$.

6. A plot of the proportion [=(Number Broken)/12] of sections of pipe broken against the weight of the striker is markedly curved, but a plot of the logit

[$=\log_e$((Number Broken+0.5)/(12.5-Number Broken)] against weight is reasonably straight.





7. (a) Exponential, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$: We have that $\mathrm{E}(Y_i) = 1/\theta$ and $\mathrm{var}(Y_i) = 1/\theta^2$, thus $\mathrm{E}(\bar{Y}) = 1/\theta$ and $\mathrm{var}(\bar{Y}) = 1/n\theta^2$.

The log likelihood, score, and derivative of the score are:

$$l(\theta) = n(\log\theta - \bar{y}\theta) \qquad U = n\left(\frac{1}{\theta} - \bar{y}\right) \qquad U' = -\frac{n}{\theta^2}$$

Thus

$$\begin{aligned}
\mathrm{E}(U) &= n\left\{\frac{1}{\theta} - \mathrm{E}(\bar{Y})\right\} = 0 \\
\mathrm{var}(U) &= \mathrm{E}(U^2) - \mathrm{E}(U)^2 = \mathrm{E}(U^2) \\
\mathrm{E}(U^2) &= n^2\mathrm{E}\left[\frac{1}{\theta} - \bar{Y}\right]^2 \\
&= n^2\mathrm{var}(\bar{Y}) \\
&= n^2\mathrm{var}(Y_i)/n = \frac{n}{\theta^2} \\
-\mathrm{E}(U') &= \frac{n}{\theta^2}
\end{aligned}$$

(b) Binomial (m known), $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$: We have that $\mathrm{E}(Y_i) = m\theta$ and $\mathrm{var}(Y_i) = m\theta(1-\theta)$, thus $\mathrm{E}(\bar{Y}) = m\theta$ and $\mathrm{var}(\bar{Y}) = m\theta(1-\theta)/n$.

The log likelihood, score, and derivative of the score are:

$$\begin{aligned}
l(\theta) &= n\left[\bar{y}\{\log\theta - \log(1-\theta)\} + m\log(1-\theta)\right] + \text{constant} \\
U &= n\left[\bar{y}\left\{\frac{1}{\theta} + \frac{1}{(1-\theta)}\right\} - \frac{m}{(1-\theta)}\right] \\
U' &= n\left[\bar{y}\left\{-\frac{1}{\theta^2} + \frac{1}{(1-\theta)^2}\right\} - \frac{m}{(1-\theta)^2}\right]
\end{aligned}$$

Thus

$$\begin{aligned}
\mathrm{E}(U) &= n\left[\mathrm{E}(\bar{Y})\left\{\frac{1}{\theta} + \frac{1}{(1-\theta)}\right\} - \frac{m}{(1-\theta)}\right] \\
&= nm\left[\theta\left\{\frac{1}{\theta(1-\theta)}\right\} - \frac{1}{(1-\theta)}\right] = 0
\end{aligned}$$

$$
\begin{aligned}
\operatorname{var}(U) &= \mathrm{E}(U^2) - \mathrm{E}(U)^2 = \mathrm{E}(U^2) \\[2ex]
\mathrm{E}(U^2) &= n^2 \mathrm{E}\left[\bar{Y}\left\{\frac{1}{\theta(1-\theta)}\right\} - \frac{m}{(1-\theta)}\right]^2 \\
&= n^2 \mathrm{E}\left(\bar{Y} - m\theta\right)^2 / \theta^2 (1-\theta)^2 \\
&= n^2 \operatorname{var}(\bar{Y}) / \theta^2 (1-\theta)^2 \\
&= n^2 \operatorname{var}(Y_i) / n\theta^2 (1-\theta)^2 \\
&= \frac{nm}{\theta(1-\theta)} \\[2ex]
-\mathrm{E}(U') &= -n\left[\mathrm{E}(\bar{Y})\left\{-\frac{1}{\theta^2} + \frac{1}{(1-\theta)^2}\right\} - \frac{m}{(1-\theta)^2}\right] \\
&= -nm\left[\theta\left\{-\frac{1}{\theta^2}\right\} - \frac{1-\theta}{(1-\theta)^2}\right] \\
&= nm\left[\frac{1}{\theta} + \frac{1}{1-\theta}\right] \\
&= \frac{nm}{\theta(1-\theta)}
\end{aligned}
$$