Example: GLM, Inference; Cell differentiation

Cell differentiation: Researchers are interested in the effect of two agents of immuno-activating ability that may introduce cell differentiation.

Outcome: cell count

Covariates: TNF (tumor necrosis factor), IFN (interferon)

(a) Write down a model which would allow the TNF effect to depend on IFN.

- Systematic component:

  $log(\lambda_i) = \beta_0 + \beta_1 TNF_i + \beta_2 IFN_i + \beta_3 TNF_i \times IFN_i$

- Random component:

$$Y_i \sim Poisson(\lambda_i)$$

(b) Estimate β and their (Wald) confidence intervals using IRWLS.

- Estimation of $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$ can be carried out using IRWLS (See the SAS code).
- 95% Wald confidence interval for $\beta_k (k = 0, \dots, 3)$ is
$$\widehat{\beta}_k \pm 1.96\widehat{SE}(\widehat{\beta}_k).$$
Since $\widehat{Var}(\widehat{\beta}) = I(\widehat{\beta})^{-1}$, $\widehat{SE}(\widehat{\beta}_k)$ is the square root of the kth diagonal element of $I(\widehat{\beta})^{-1}$. For the Poisson regression,
$$I(\widehat{\beta}) = X'VX,$$
where $V = diag\{v(\mu_1), ..., v(\mu_n)\}$ with $v(\mu) = \mu$.

(c) Test for IFN and interactions between TNF and IFN. Specifically, write out the Wald statistic and compute it using IML.

- H_0: $\beta_2 = \beta_3 = 0$ vs H_1: $\beta_2 \neq 0$ or $\beta_3 \neq 0$

- Contrast matrix

$$C = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Under H_0, $C\widehat{\beta}$ asymptotically follows the multivariate normal distribution with mean zero and variance $CI(\widehat{\beta})^{-1}C^T$.

- Wald test statistic is

$$X_w^2 = \{C\widehat{\beta}\}^T\{CI(\widehat{\beta})^{-1}C^T\}^{-1}\{C\widehat{\beta}\},$$

which follows χ_2^2 under the null hypothesis.

- In this data, $X_w^2 = 4.4615902$ and the corresponding p-value=0.107443. We cannot reject the null hypothesis.

(d) This time, carry out the Score test. Write out the formula, then do the computations using IML.

- H_0: $\beta_2 = \beta_3 = 0$ vs H_1: $\beta_2 \neq 0$ or $\beta_3 \neq 0$
- Let $\widehat{\beta}_H = (\widehat{\beta}_{0H}, \widehat{\beta}_{1H}, 0, 0)$ be the MLE of β under

the null hypothesis. Score test statistic is

$$X_s^2 = U(\widehat{\beta}_H)^T I(\widehat{\beta}_H)^{-1} U(\widehat{\beta}_H),$$

which follows χ_2^2 under the null hypothesis.

- In this data, $X_s^2 = 4.4782374$ and the corresponding p-value=0.1065524. We cannot reject the null hypothesis.
- [Here I used the same notation in the note ($\widehat{\beta}_H$). In the today's lecture I used $\widehat{\beta}^0$ instead of $\widehat{\beta}_H$. But they are the same.]

(e) Carry out the likelihood ratio test (LRT). Write out the formula, then do the computations using IML.

- H_0: $\beta_2 = \beta_3 = 0$ vs H_1: $\beta_2 \neq 0$ or $\beta_3 \neq 0$
- Maximum likelihood under the null and alternative:

$$l(\widehat{\beta}_H) = \sum_{i=1}^{n} \{Y_i \log(\widehat{\lambda}_{H,i}) - \widehat{\lambda}_{H,i}\}$$

$$l(\widehat{\beta}) = \sum_{i=1}^{n} \{Y_i \log(\widehat{\lambda}_i) - \widehat{\lambda}_i\}$$

where $\widehat{\lambda}_{H,i} = exp(X_i^T \widehat{\beta}_H)$ and $\widehat{\lambda}_i = exp(X_i^T \widehat{\beta})$.

- Likelihood ratio test statistic is

$$X_L^2 = 2\{l(\widehat{\beta}) - l(\widehat{\beta}_H)\},$$

which follows χ_2^2 under the null hypothesis.

- In this data, $X_L^2 = 4.3365661$ and the corresponding p-value=0.1143738. We cannot reject the null hypothesis.
- [In the today's lecture I used $\widehat{\lambda}^0$ instead of $\widehat{\lambda}_H$. But they are the same.]

(g) Carry out the LRT using proc genmod by fitting full and reduced models.

- From the proc genmod, $l(\widehat{\beta}_H) = 1208.7574$ and $l(\widehat{\beta}) = 1210.9257$. LRT test statistic is

$$X_L^2 = 2 * \{l(\widehat{\beta}) - l(\widehat{\beta}_H)\} = 4.3366$$

- P-value is 0.1143.

(h) Recompute the Wald and LRT, this time using the Contrast statement.

- See the SAS code.