Nr. of boys	0	1	2	3	4	5	6	7	8	9	10	11	12
Nr. of families	3	24	104	286	670	1033	1343	1112	829	478	181	45	7

Example 21. The following data gives the number of male children among the first 12 children of family size 13 in 6115 families taken from hospital records in the 19th century Saxony. The 13th child is ignored to assuage the effect of families non-randomly stopping when a desired gender is reached. Test the null hypothesis that the gender of the babies can be viewed as realisations of independent random variables having the same probability of a baby boy for each family.

Hint. Let X_i stand for the number of boys in the i-th family ($i \in \{1, \ldots, n\}$, where n stands for the sample size). Then the counts in the table can be represented by

$$n_k = \sum_{i=1}^{n} \mathbb{I}\{X_i = k\}, \quad k = 0, 1, \ldots, 12$$

and the table can be viewed as a realisation of a random vector $(n_0, n_1, \ldots, n_{12})^{\mathsf{T}}$ that follows multinomial distribution $\mathsf{Mult}_{13}(n, \boldsymbol{\pi})$.

Note that under the null hypothesis X_i follows the binomial distribution, thus

$$\pi_k = \mathsf{P}(X_i = k) = \binom{12}{k} p^k (1-p)^{12-k}, \quad k = 0, 1, \ldots, 12,$$

where $p \in (0, 1)$ is the probability of baby boy.

Thus to parametrize the problem (so that it fits into the framework of this section) put $\psi = p$ and get

$$\pi_0 = (1-\psi)^{12}, \quad \pi_k = \binom{n}{k} \psi^k (1-\psi)^{12-k} + \tau_k, \quad k = 1, \ldots, 11,$$

and $\pi_{12} = 1 - \sum_{k=0}^{11} \pi_k$. The hypotheses can now be written as

$$H_0 : (\tau_1, \ldots, \tau_{11})^{\mathsf{T}} = \mathbf{0}_{11}, \qquad H_1 : (\tau_1, \ldots, \tau_{11})^{\mathsf{T}} \neq \mathbf{0}_{11}.$$

Nevertheless it would take some time to derive either the Wald statistic (W_n^*) or Rao score statistic (R_n^*) as one needs to calculate the score statistic and (empirical) Fisher information matrix.

On the other hand using (29) it is straightforward to calculate the likelihood ratio test LR_n^* as

$$\sup_{\boldsymbol{\theta} \in \Theta} \ell_n(\boldsymbol{\theta}) = \sum_{k=0}^{12} n_k \log\left(\tfrac{n_k}{n}\right)$$

and

$$\sup_{\boldsymbol{\theta}\in\Theta_0} \ell_n(\boldsymbol{\theta}) = \sum_{k=0}^{12} n_k \log \widetilde{\pi}_k, \quad \text{where} \quad \widetilde{\pi}_k = \binom{12}{k} \big(\widetilde{\psi}_n\big)^k \big(1-\widetilde{\psi}_n\big)^{12-k}, \quad \text{with} \quad \widetilde{\psi}_n = \sum_{k=1}^{12} \tfrac{k\, n_k}{12\, n}.$$

By Theorem 7 the test statistic LR_n^* converges under the null hypothesis to χ^2-distribution with 11 degrees of freedom.

Another approach to test the hypothesis of interest would be (to forget about the test statistics LR_n^*, W_n^*, R_n^* and) to use the standard χ^2-test of goodness-of-fit in multinomial distribution with estimated parameters. The test statistics would be

$$X^2 = \sum_{k=0}^{12} \frac{(n_k - n\,\widetilde{\pi}_k)^2}{n\,\widetilde{\pi}_k} \tag{30}$$

and under the null hypothesis it has also asymptotically χ^2-distribution with 11 degrees of freedom. In fact it can be proved* that the test statistic X^2 given by (30) corresponds to the test statistic of the Rao score test (R_n^*) with $I^{11}(\widetilde{\boldsymbol{\theta}}_n)$ taken as $\widehat{I}_n^{11}$.

Example 22. Breusch-Pagan test of heteroscedasticity.

Example 23. Suppose that you observe independent identically distributed random vectors $(\boldsymbol{X}_1^\mathsf{T}, Y_1)^\mathsf{T}, \ldots, (\boldsymbol{X}_n^\mathsf{T}, Y_n)^\mathsf{T}$ such that

$$\mathsf{P}(Y_1 = 1 \,|\, \boldsymbol{X}_1) = \frac{\exp\{\alpha + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{X}_1\}}{1+\exp\{\alpha + \boldsymbol{X}^\mathsf{T}\boldsymbol{\beta}_1\}}, \qquad \mathsf{P}(Y_1 = 0 \,|\, \boldsymbol{X}_1) = \frac{1}{1+\exp\{\alpha + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{X}_1\}},$$

where the distribution of $\boldsymbol{X}_1 = (X_{11}, \ldots, X_{1d})^\mathsf{T}$ does not depend on the unknown parameters α a $\boldsymbol{\beta}$.

(i) Derive a test for the null hypothesis $H_0 : \boldsymbol{\beta} = \boldsymbol{0}_d$ against the alternative that $H_1 : \boldsymbol{\beta} \neq \boldsymbol{0}_d$.

(ii) Find the confidence set for the parameter $\boldsymbol{\beta}$.

Literature: Anděl (2007) Chapter 8.6, Kulich (2014), Zvára (2008) pp. 122–128.

2.7 Profile likelihood†

Let $\boldsymbol{\theta}$ be divided into $\boldsymbol{\tau}$ containing the first q components ($1 \leq q < p$) and $\boldsymbol{\psi}$ containing the remaining $p-q$ components, i.e.

$$\boldsymbol{\theta} = (\boldsymbol{\tau}^\mathsf{T}, \boldsymbol{\psi}^\mathsf{T})^\mathsf{T} = (\theta_1, \ldots, \theta_q, \theta_{q+1}, \ldots, \theta_p)^\mathsf{T}.$$

* More precisely, it is said so in the textbooks but I have not managed to find the derivation. † *Profilová věrohodnost.*

Ex 22) „TEST OF HETEROSCEDASTICITY"

$\begin{pmatrix} Y_1 \\ X_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ X_n \end{pmatrix}$ i.i.d

$Y_i = X_i^T \beta + \exp(\lambda^T z_i)\, \varepsilon_i$, $\varepsilon_i \perp X_i$, $var(\varepsilon_i) = \sigma^2$

$\uparrow$
$z_i = g(X_i)$

$H_0 : \lambda = 0 \qquad H_1 : \lambda \neq 0$

BREUSCH - PAGAN TEST

- ASSUME THAT $\varepsilon_i \sim N(0, \sigma^2)$
- TEST STAT. : $R_n^* = \frac{1}{n} U_{1n}^T(\tilde{\theta}_n)\, I_n^{11}\, U_{1n}(\tilde{\theta}_n)$,

$$\theta = \begin{pmatrix} \boxed{\lambda} \\ \boxed{\begin{matrix} \sigma^2 \\ \beta \end{matrix}} \end{pmatrix} \quad \begin{matrix} \leftarrow \xi \\ \leftarrow \psi \end{matrix}$$

DENOTE $h_i(\lambda) = \exp(\lambda^T z_i)$

$\leadsto$ note that $Y_i \mid X_i \sim N(X_i^T \beta, \sigma^2 h_i^2(\lambda))$

$\leadsto$ $$l_n(\theta) = \log\left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi h_i^2(\lambda)}} \cdot \exp\left\{ -\frac{1}{2\sigma^2 h_i^2(\lambda)} \sum_{i=1}^n (Y_i - X_i^T \beta)^2 \right\} \right) + c$$

$$= -\frac{n}{2} \log \sigma^2 - \sum_{i=1}^n \log h_i(\lambda) - \frac{1}{2\sigma^2} \sum_{i=1}^n \frac{(Y_i - X_i^T \beta)^2}{h_i^2(\lambda)} + c$$

$$= -\frac{n}{2} \log \sigma^2 - \sum_{i=1}^n \lambda^T z_i - \frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - X_i^T \beta) \exp(-2\lambda^T z_i) + c$$

UNDER $H_0: \underset{\sim}{\lambda} = \underset{\sim}{0}$

$$\rightarrow \tilde{\ell}_n(\sigma^2, \underset{\sim}{\beta}) = \ell_n(\underset{\sim}{0}, \sigma^2, \underset{\sim}{\beta}) =$$

$$= -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta})^2$$

$$\rightarrow \underset{\sim}{\tilde{\beta}}_n = \left(\sum_{i=1}^{n} \underset{\sim}{X}_i \underset{\sim}{X}_i^T \right)^{-1} \sum_{i=1}^{n} \underset{\sim}{X}_i Y_i =: \underset{\sim}{\hat{\beta}}_{LS}$$

$$\tilde{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\hat{\beta}}_{LS})^2$$

$$\rightarrow \underset{\sim}{\tilde{\theta}}_n = \begin{pmatrix} \underset{\sim}{0} \\ \tilde{\sigma}_n^2 \\ \underset{\sim}{\tilde{\beta}}_n \end{pmatrix}$$

$$\underset{\sim}{U}_{1n}(\underset{\sim}{\theta}) = \frac{\partial \ell_n(\underset{\sim}{\theta})}{\partial \underset{\sim}{\lambda}} = -\sum_{i=1}^{n} \underset{\sim}{Z}_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta})^2 \exp(-2 \underset{\sim}{\lambda}^T \underset{\sim}{Z}_i) \cdot (-2 \underset{\sim}{Z}_i)$$

$$\rightarrow \underset{\sim}{U}_{1n}(\underset{\sim}{\tilde{\theta}}_n) = -\sum_{i=1}^{n} \underset{\sim}{Z}_i + \frac{1}{\tilde{\sigma}_n^2} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\tilde{\beta}}_n)^2 \underset{\sim}{Z}_i$$

$$= -\sum_{i=1}^{n} \underset{\sim}{Z}_i + \frac{1}{\tilde{\sigma}_n^2} \sum_{i=1}^{n} \tilde{\varepsilon}_i^2 \underset{\sim}{Z}_i$$

$$= -n \underset{\sim}{\bar{Z}}_n \cdot \frac{1}{\tilde{\sigma}_n^2} \frac{1}{n} \sum_{i=1}^{n} \tilde{\varepsilon}_i^2 + \frac{1}{\tilde{\sigma}_n^2} \sum_{i=1}^{n} \tilde{\varepsilon}_i^2 \underset{\sim}{Z}_i$$

$$= \frac{1}{\tilde{\sigma}_n^2} \sum_{i=1}^{n} \tilde{\varepsilon}_i^2 \cdot (\underset{\sim}{Z}_i - \underset{\sim}{\bar{Z}}_n)$$

FURTHER WE USE $I_{11.2}(\underset{\sim}{\tilde{\theta}}_n)$ AS ESTIMATE OF I_n^{11}

$$\frac{\partial \ell_n(\underset{\sim}{\theta})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta})^2 h_i^2(\underset{\sim}{\lambda})$$

$$\frac{\partial \ell_n(\underset{\sim}{\theta})}{\partial \underset{\sim}{\beta}} = + \frac{1}{\sigma^2} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta}) \underset{\sim}{X}_i h_i^2(\underset{\sim}{\lambda})$$

FOR $\underset{\sim}{\lambda} = \underset{\sim}{0}$

$$\frac{\partial^2 \ell_n(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} = \begin{pmatrix} -\frac{2}{\sigma^2} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta})^2 \underset{\sim}{Z}_i \underset{\sim}{Z}_i^T, & -\frac{1}{\sigma^4} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta})^2 \underset{\sim}{Z}_i, & -\frac{2}{\sigma^2} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta}) \underset{\sim}{Z}_i \underset{\sim}{X}_i^T \\ -\frac{1}{\sigma^4} \sum (\dots)^2 \underset{\sim}{Z}_i^T, & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta})^2, & -\frac{1}{\sigma^4} \sum_{i=1}^{n} (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta}) \underset{\sim}{X}_i \\ -\frac{2}{\sigma^4} \sum (\dots) \underset{\sim}{Z}_i^T \underset{\sim}{X}_i, & -\frac{1}{\sigma^4} \sum (\dots) \underset{\sim}{X}_i^T, & -\frac{1}{\sigma^2} \sum_{i=1}^{n} \underset{\sim}{X}_i \underset{\sim}{X}_i^T \end{pmatrix}$$

$$\rightarrow I(\underset{\sim}{\theta}) = \begin{pmatrix} 2 \mathbb{E} \underset{\sim}{Z}_i \underset{\sim}{Z}_i^T, & \frac{1}{\sigma^2} \mathbb{E} \underset{\sim}{Z}_i, & \underset{\sim}{0} \\ \frac{1}{\sigma^2} \mathbb{E} \underset{\sim}{Z}_i^T, & \frac{1}{2\sigma^4}, & \underset{\sim}{0} \\ \underset{\sim}{0}, & \underline{0}, & \frac{1}{\sigma^2} \mathbb{E} \underset{\sim}{X}_i \underset{\sim}{X}_i^T \end{pmatrix}$$

$\varepsilon_i = (Y_i - \underset{\sim}{X}_i^T \underset{\sim}{\beta}) \perp \underset{\sim}{X}_i$

$$\rightarrow I_{11.2}(\underset{\sim}{\theta}) = 2 \mathbb{E} \underset{\sim}{Z}_i \underset{\sim}{Z}_i^T - \frac{1}{\sigma^2} \mathbb{E} \underset{\sim}{Z}_i \cdot 2\sigma^4 \cdot \frac{1}{\sigma^2} \mathbb{E} \underset{\sim}{Z}_i^T =$$

$$= 2 \mathbb{E} \left\{ (\underset{\sim}{Z}_i - \mathbb{E} \underset{\sim}{Z}_i)(\underset{\sim}{Z}_i - \mathbb{E} \underset{\sim}{Z}_i)^T \right\} = 2 \operatorname{var}(\underset{\sim}{Z}_i)$$

$$\rightarrow \widehat{I}_{11.2} = \frac{2}{n} \sum_{i=1}^{n} (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n)(\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n)^T =: 2 \, \underset{\sim}{S}_Z$$

$$\rightarrow R_n^* = \underbrace{\frac{1}{2n} \sum_{i=1}^{n} \frac{\hat{\varepsilon}_i^2}{\hat{\sigma}_n^2} (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n)^T}_{\text{EMPIRIC. COVAR. OF } \frac{\hat{\varepsilon}_i^2}{\hat{\sigma}_n^2} \text{ AND } \underset{\sim}{Z}_i} \underset{\sim}{S}_Z^{-1} \sum_{i=1}^{n} \frac{\hat{\varepsilon}_i^2}{\hat{\sigma}_n^2} (\underset{\sim}{Z}_i - \bar{\underset{\sim}{Z}}_n)^T$$

KOENKER: 2 $\$_2$ ESTIMATES

$$\mathrm{var}_{H_0}\left(\frac{\varepsilon_i^2}{\sigma^2}(z_i - \mathbb{E} z_i)\right) = \ldots = \mathrm{var}\left(\frac{\varepsilon_i^2}{\sigma^2}\right)\mathrm{var}(z_i)$$

FOR $\varepsilon_i \sim N(0,\sigma^2)$

$$\mathrm{var}\left(\frac{\varepsilon_i^2}{\sigma^2}\right) = \mathrm{var}\left(\text{"}N(0,1)^{\text{"}}\right) = 2$$

FOR $\varepsilon_i \not\sim N(0,\sigma^2)$ $\rightsquigarrow$ ONE CAN ESTIMATE

$$\mathrm{var}\left(\frac{\varepsilon_i^2}{\sigma^2}\right) = \frac{1}{\sigma^4}\mathbb{E}(\varepsilon_i^2 - \sigma^2)^2 \quad \text{BY}$$

$$\tilde{\gamma}_n = \frac{1}{\tilde{\sigma}_n^4}\frac{1}{n}\sum_{i=1}^{n}(\tilde{\varepsilon}_i^2 - \tilde{\sigma}^2)^2$$

$$\rightarrow T_n^{(k)} = \frac{1}{\tilde{\gamma}_n n}\sum_{i=1}^{n}\frac{\tilde{\varepsilon}_i^2}{\tilde{\sigma}_n^2}(z_i - \bar{z}_n)^T \$_z^{-1}\sum_{i=1}^{n}\frac{\tilde{\varepsilon}_i^2}{\tilde{\sigma}_n^2}(z_i - \bar{z}_n)$$

IN PRACTICE: $\tilde{\varepsilon}_i^2 \sim \beta_0 + \beta^T z_i$

$$H_0: \beta = 0$$

Ex 25) RVT $Y^{(\lambda)} \begin{cases} \frac{Y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log Y, & \lambda = 0 \end{cases}$

ASSUMED MODEL:

$$Y_i^{(\lambda)} = X_i^T \beta + \varepsilon_i \quad i = 1, \dots, n$$
$$\varepsilon_i \sim N(0, \sigma^2)$$

WE NEED $f_{Y_i | X_i}$

$$F_{Y_i|X_i}(y) = P(Y_i \leq y | X_i) = P(Y_i^{(\lambda)} \leq y^{(\lambda)} | X_i) = F_{Y_i^{(\lambda)}|X_i}(y^{(\lambda)})$$

$$\leadsto f_{Y_i|X_i}(y) = F'_{Y_i|X_i}(y) = f_{Y_i^{(\lambda)}|X_i}(y^{(\lambda)})\, y^{\lambda-1}$$

$$\leadsto \ell_n(\lambda, \beta, \sigma^2) = \log\left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{(Y_i^{(\lambda)} - X_i^T \beta)^2}{2\sigma^2} \right\} Y_i^{\lambda-1} \right) + c$$

$$= -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i^{(\lambda)} - X_i^T \beta)^2 + (\lambda - 1) \log Y_i + c$$

FOR FIXED λ:

$$\frac{\partial \ell_n(\lambda, \beta, \sigma^2)}{\partial \beta} \stackrel{!}{=} 0 \quad \leadsto \tilde{\beta}_n(\lambda) = \left(\sum_{i=1}^n X_i X_i^T \right)^{-1} \left(\sum_{i=1}^n X_i Y_i^{(\lambda)} \right)$$

$$\frac{\partial \ell_n(\lambda, \beta, \sigma^2)}{\partial \sigma^2} \stackrel{!}{=} 0 \quad \tilde{\sigma}_n^2(\lambda) = \frac{1}{n} \sum_{i=1}^n (Y_i^{(\lambda)} - X_i^T \tilde{\beta}_n(\lambda))^2$$

$$\ell_n^{(p)}(\lambda) = \max_{\beta, \sigma^2} \ell_n(\lambda, \beta, \sigma^2) =$$
$$= -\frac{n}{2} \log \tilde{\sigma}_n^2(\lambda) - \frac{1}{2\tilde{\sigma}_n^2(\lambda)} \sum_{i=1}^n (Y_i^{(\lambda)} - X_i^T \tilde{\beta}_n(\lambda))^2 + (\lambda - 1) \sum_{i=1}^n \log Y_i + c$$

$$= -\frac{n}{2} \log \tilde{\sigma}_n^2(\lambda) - \frac{n}{2} + (\lambda - 1) \sum_{i=1}^n \log Y_i + c$$

CI for λ: $\left\{ \lambda : 2(\ell_n^{(p)}(\hat{\lambda}_n) - \ell_n^{(p)}(\lambda)) \leq \chi_1^2(1-\alpha) \right\}$, where $\hat{\lambda}_n = \underset{\lambda > 0}{\operatorname{argmax}}\, \ell_n^{(p)}(\lambda)$.

BOXCOX (R-ENVIRONMENT, MASS)

$$Y_i^* = \frac{Y_i}{\exp\{\frac{1}{n}\sum_{i=1}^n \log Y_i\}} \leadsto \sum_{i=1}^n \log Y_i^* = 0$$

$$\leadsto l_n^{(n)}(\lambda) = -\frac{n}{2}\log \sum_{i=1}^n \left(Y_i^{*(\lambda)} - \underset{\sim}{X}_i^T \underset{\sim}{\beta}_n^*(\lambda)\right)^2,$$

$$\text{where } \underset{\sim}{\beta}_n^*(\lambda) = \left(\sum_{i=1}^n \underset{\sim}{X}_i \underset{\sim}{X}_i^T\right)^{-1}\left(\sum_{i=1}^n \underset{\sim}{X}_i^T Y_i^{*(\lambda)}\right)$$