

第六章 偏大离差的统计推断

Overdispersion

- Often the data we observe is more variable than we can handle using the GLM models considered so far.

- Defn: Overdispersion occurs for GLMs when

$$\text{var}(Y_i) > V(\mu_i)a(\phi_i).$$

- Examples:

1. For $Y_i \sim B(m_i, p_i)$ (with $m_i > 1$) we have

$$\begin{aligned} E(Y_i) &= m_i p_i = \mu_i; \\ \text{var}(Y_i) &> m_i p_i (1 - p_i) = \frac{\mu_i(m_i - \mu_i)}{m_i}. \end{aligned}$$

2. For $Y_i \sim Po(\mu_i)$ we have

$$\begin{aligned} E(Y_i) &= \mu_i; \\ \text{var}(Y_i) &> \mu_i. \end{aligned}$$

The implications of overdispersion

- In a GLM with overdispersion, the standard inferences do not apply because, we underestimate standard errors and thus the tests are anti-conservative.

An example

- Let us re-consider the beetle dataset;
Remember we have the number of beetles that died after five hours exposure to a different number of concentrations of carbon disulfide(doses).The concentration units are $\log_{10} \text{CS}_2 \text{mg/L}$.
- Suppose that Y_1, \dots, Y_8 are independent $B(m_i, p_i)$ RVs where for each i:
 - m_i are the number of beetles exposed at certain dose combination x_i ;
 - we model p_i by

$$\log\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta x_i.$$

Fitting the model in R

Call:

```
glm(formula = prop ~ dose, family = binomial, data = beetles,  
     weights = number)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5941	-0.3944	0.8329	1.2592	1.5940

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-60.717	5.181	-11.72	<2e-16 ***
dose	34.270	2.912	11.77	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 284.202 on 7 degrees of freedom
Residual deviance: 11.232 on 6 degrees of freedom
AIC: 41.43

Number of Fisher Scoring iterations: 4

Analysis of deviance table

Analysis of Deviance Table

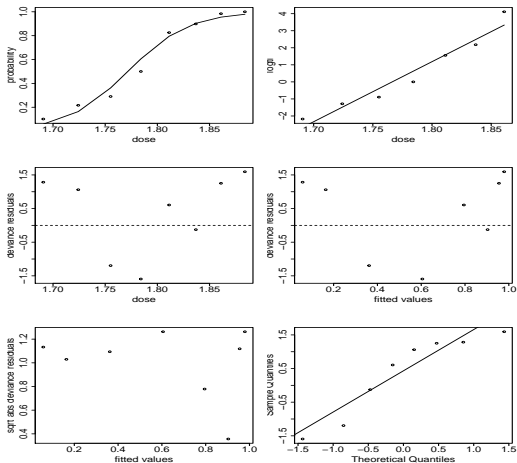
Model: binomial, link: logit

Response: prop

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev
NULL			7	284.202
dose 1	1	272.97	6	11.232

Diagnostic plots



Identifying Overdispersion

- Variance of deviance residuals are not very constant when we plot the fitted values versus the deviance residuals (or the square root of absolute deviance residuals)
- normal Q-Q plot of deviance residuals indicate S shape - deviance residuals are not normal. There are not enough residuals near zero. The data may be more variable than can be explained using a binomial distribution
- Are we sure?(after all, we only have eight deviance residuals)

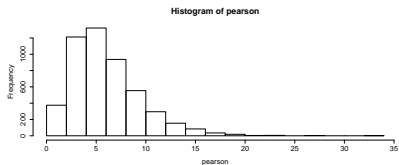
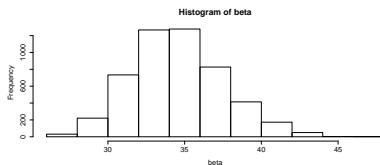
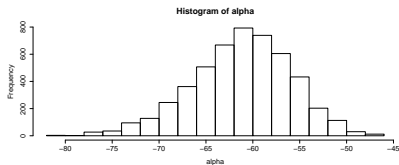
A simulation study

- We will simulate binomial data that are not overdispersed.
- Let $\hat{\alpha}$ and $\hat{\beta}$ denote the estimates from our logistic regression model.
- Here is our simulation procedure;
 1. Simulate $Y_i^* \sim B(m_i, p_i^*)$, $i = 1, \dots, 8$ with
$$\log\left(\frac{p_i^*}{1-p_i^*}\right) = \hat{\alpha} + \hat{\beta}x_i,$$
 2. Fit the binomial GLM model to there data.
 3. Repeat.
- We will summarize the estimates, diagnostic statistics and residual plots.

Simulation results:the coefficients

- based on 5000 replications –Estimated ES for $\hat{\alpha}$ is 5.325
(95% bootstrap interval is [5.217,5.431])
–Estimated ES for $\hat{\beta}$ is 2.993
(95% bootstrap interval is [2.935,3.054])
- Compare with the values for our data;
–Estimated ES for $\hat{\alpha}$ is 5.181.
–Estimated ES for $\hat{\beta}$ is 2.912
- We slightly underestimate the SEs in our model.

Distribution of the statistics

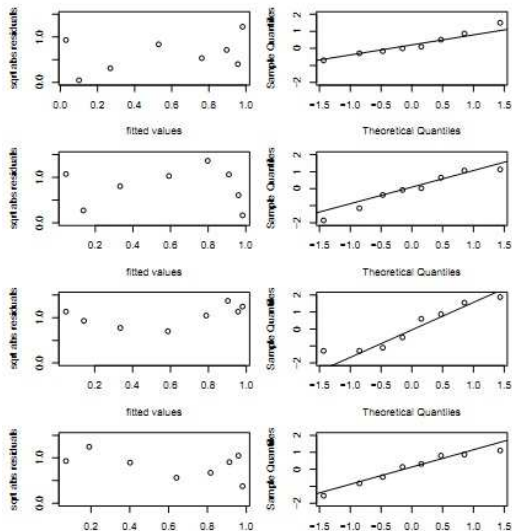


The Pearson χ^2 Statistic

- Our observed χ^2 statistic is 10.027.
 - We calculate this value using the fact that χ^2 =sum of squares of Pearson residuals.
- From the simulations,the mean χ^2 statistics value is 5.950,and the variance is 11.63.
 - which agrees closely with homework 2!

Residual plots and normal Q-Q plots

Plots for four randomly chosen simulations:



Modeling dispersion through a scale parameter.

- For both the binomial and Poisson GLMs, $a(\phi) = 1$
- One way is to introduce a dispersion parameter $\sigma^2 > 0$, and let

$$a(\phi) = \sigma^2.$$

- Then, $\text{var}(Y_i) = a(\phi)V(\mu_i) = \sigma^2 V(\mu_i)$.
- Why use this model? Think about the case for a normal GLM.
- Introducing this parameter often causes the pmf/pdf for Y_i to no longer be an exponential family.

Estimating σ^2

- Rather than introducing σ^2 into the likelihood equations, McCullagh and Nelder suggest the estimator

$$\hat{\sigma}^2 = \frac{\chi^2}{n-p}$$

- We estimate β using the same IWLS algorithm.
- Under the true model, $\hat{\sigma}^2$ is a consistent estimate of σ^2 for increasing n (and fixed m_i in the binomial case).
- For our observed data $n = 8, p = 2$, and $\chi^2 = 10.027$. Hence

$$\hat{\sigma}^2 = \frac{10.027}{8-2} = 1.67.$$

- Conclusion:

Adjusting the inference for overdispersion

- Here is the coefficient table for our model (assuming no overdispersion):

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-60.717	5.181	-11.72	< 2e - 16 * **
dose	34.270	2.912	11.77	< 2e - 16 * **

- Using the Pearson χ^2 statistic we estimated the overdispersion parameter, σ^2 , by $\hat{\sigma}^2 = 1.67$.
- We need to adjust the standard errors of β
- We will confirm this later using quasi-likelihood but for the moment we state the result. Asymptotically

$$\hat{\beta} \sim N_p(\beta, \sigma^2(X^T W X)^{-1})$$

- We replace σ^2 by our consistent estimator above.

The adjusted coefficient table

```
## calculate the Pearson statistic
```

```
X2 <- sum(resid(model, type = "pearson")^2)/(8 - 2)
```

```
X2
```

```
[1] 1.671136
```

```
##obtain the 'naive' coefficient table
```

```
ctable <- summary(model)$coef
```

```
ctable
```

	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>Pr(> z)</i>
(Intercept)	-60.71745	5.180701	-11.71993	1.007549e - 31
dose	34.27033	2.912134	11.76811	5.698445e - 32

```
##multiple the standard errors by the estimate of 'sigma'
```

```
ctable[,2] <- ctable[,2] * sqrt(X2)
```

```
##recalculate the Z statistic
```

```
ctable[,3] <- ctable[,1] / ctable[,2]
```

```
##recalculate the P-values
```

```
ctable[,4] <- 2 * pnorm(-abs(ctable[,3]))
```

```
##display the coefficient table, adjusted for overdispersion
```

```
ctable
```

	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>Pr(> z)</i>
(Intercept)	-60.71745	6.697218	-9.066071	1.233865e - 19
dose	34.27033	3.764587	9.103343	8.759298e - 20

Interpreting the new coefficient table

- What effect does the adjustment for overdispersion have on the coefficient summary?

The analysis of deviance for overdispersion

- Suppose we fit a model with q parameters - let the estimate of μ be $\hat{\mu}^{(1)}$.

The **deviance** for this model is

$$D(y, \hat{\mu}^{(1)}) = 2\sigma^2[l(y) - l(\hat{\mu}^{(1)})]$$

- Now add $p - q$ parameter to this model. Let the estimate of μ now be $\hat{\mu}^{(2)}$.

The **deviance** for this model is

$$D(y, \hat{\mu}^{(2)}) = 2\sigma^2[l(y) - l(\hat{\mu}^{(2)})]$$

- The differences of deviances is

$$D(y, \hat{\mu}^{(1)}) - D(y, \hat{\mu}^{(2)}) = 2\sigma^2[l(\hat{\mu}^{(2)}) - l(\hat{\mu}^{(1)})]$$

Asymptotically, this difference has $\sigma^2\chi_{p-q}^2$ distribution (again we show this later!)

Adjusting the analysis of deviance table

- For our model we have

Analysis of Deviance Table

Terms added sequentially (first to last)

	<i>Df</i>	<i>Deviance</i>	<i>Resid.Df</i>	<i>Resid.Dev</i>
<i>NULL</i>			7	284.202
<i>dose</i>	1	272.970	6	11.232

- Without overdispersion** the P-value is

```
1 - pchisq(272.970, 1)
```

```
[1] 0
```

- With overdispersion** we need to divide by our estimate of σ^2 , and the P-value is

```
1 - pchisq(272.970/X2, 1)
```

```
[1] 0
```

- Conclusions:

Mechanisms for overdispersion



- The most common mechanism for overdispersion is due to **cluster** effects.
- Example clusters in a population: families, households, litters, colonies, neighborhoods.
- We expect different clusters to behave differently in the population.
- However, we often **observe** a quantity which is aggregated over these clusters.
- Exercise: what do we observe in the beetle dataset?

What are the possible clusters?

Binomial overdispersion

- **The clusters:** Suppose we have k individuals each sampled from l clusters in the population.

We observe $m = kl$ individuals in total.

- **The cluster model:** Let Z_j denote the number of successes in cluster j . Assume that $\{Z_j : j = 1, \dots, l\}$ are a set of independent RVs with $Z_j \sim B(k, p_j)$ for each j .
- **The cluster effects:** $\{p_j : j = 1, \dots, l\}$ are iid RVs with $E(p_j) = p$ and $\text{var}(p_j) = \tau^2 p(1 - p)$, for some constants $0 < p < 1$ and τ^2 .
- **What we observe:** Let $Y = \sum_{j=1}^l Z_j$.
- What are the moments of Y ?

The expectation of Y



The variance of Y



The variance of Y (cont.)



A parametric model for Binomial overdispersion

(Crowder, 1978; Plackett, 1981; Williams, 1982)

- Consider the following hierarchical model for Y :

$$Y|p \sim B(m,p);$$

$$p \sim \text{Beta}(\alpha, \beta).$$

- Then it can be shown that Y has **Beta-Binomial** distribution with parameters m , α and β .

(For more details see homework 3).

- The moments of Y are

$$E(Y) = m\pi;$$

$$\text{var}(Y) = m\pi(1-\pi)[1 + (m-1)\tau^2].$$

where π and τ^2 are expressed in terms of α and β .

Modeling using the beta-binomial

- Method 1: Ignore the true distribution. Fit a binomial GLM with overdispersion parameter, $\sigma^2 (= [1 + (m - 1)\tau^2])$
- Method 2: Ignore the true distribution. Fit a binomial GLM but use an amended variance term which allows us to estimate τ^2 .
(Need a different IWLS algorithm - see later).

Modeling using the beta-binomial

- We have that

$$E(Y) =$$

from which we deduce that

$$\pi =$$

- Method3: model parametrically using the beta-binomial:
- We estimate the parameters via maximum likelihood.

Beta-Binomial分布

- 求解Beta-Binomial分布的概率质量函数，同时推导其期望方差。
- 已知：

$$Y|p \sim B(m, p), p \sim \text{Beta}(\alpha, \beta)$$

$$\pi = E(p) = \frac{\alpha}{\alpha + \beta}, \text{Var}(p) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

求Y的无条件概率质量函数。

Beta-Binomial分布质量函数推导过程

- 依题意可得

$$f(y|p) = \binom{m}{y} p^y (1-p)^{m-y} \quad f(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}$$

- 由乘法公式可得 y, p 的联合概率密度函数为：

$$f(y, p) = f(p)f(y|p) = \binom{m}{y} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha+y-1} (1-p)^{m+\beta-y-1}$$

Beta-Binomial分布质量函数推导过程

- 从而可以求得 y 的边缘分布函数为：

$$\begin{aligned} f(y) &= \int_0^1 \binom{m}{y} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha+y-1} (1-p)^{m+\beta-y-1} dp \\ &= \binom{m}{y} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + y)\Gamma(m + \beta - y)}{\Gamma(\alpha + \beta + m)} \\ &\quad \int_0^1 \frac{\Gamma(\alpha + \beta + m)}{\Gamma(\alpha + y)\Gamma(m + \beta - y)} p^{\alpha+y-1} (1-p)^{m+\beta-y-1} dp \\ &= \binom{m}{y} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + y)\Gamma(m + \beta - y)}{\Gamma(\alpha + \beta + m)} \end{aligned}$$

- 记为： $Y \sim BB(\alpha, \beta, m)$

Beta-Binomial分布的期望和方差

- 下面求Y的期望与方差（利用重期望和重方差公式）：

$$E(Y) = E_p[E(Y|p)] = E_p[mp] = m\pi$$

$$\begin{aligned} \text{Var}(Y) &= E_p[\text{Var}(Y|p)] + \text{Var}(E_p(Y|p)) \\ &= E_p[mp(1-p)] + \text{Var}(mp) \\ &= m(\pi - \text{Var}(p) - \pi^2) + m^2 \text{Var}(p) \\ &= m\left(\pi - \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} - \pi^2\right) \\ &\quad + m^2 \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= m\pi(1-\pi) \left(\frac{\alpha+\beta+m}{\alpha+\beta+1}\right) \end{aligned}$$

Beta-Binomial分布的期望和方差的另一种求法

- 也可由Beta-Binomial分布的概率分布直接计算其期望和方差。

$$\begin{aligned} E(Y) &= \sum_{y=0}^{\infty} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + y)\Gamma(m + \beta - y)}{\Gamma(\alpha + \beta + m)} \frac{m!}{(m - y)!y!} y \\ &= \sum_{y=1}^{\infty} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma[\alpha + 1 + (y - 1)]\Gamma[\beta + (m - 1) - (y - 1)]}{\Gamma[(\alpha + 1) + \beta + (m - 1)]} \\ &\quad \cdot \frac{m(m - 1)!}{[m - 1 - (y - 1)]!(y - 1)!} \\ &= \text{令 } k = y - 1 m \sum_{y=0}^{\infty} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + 1 + k)\Gamma[\beta + (m - 1) - k]}{\Gamma[\alpha + 1 + \beta + (m - 1)]} \\ &\quad \cdot \frac{(m - 1)!}{(m - 1 - k)!k!} \end{aligned}$$

Beta-Binomial分布的期望和方差的另一种求法

$$\begin{aligned} &= \frac{m\alpha}{\alpha + \beta} \\ &\quad \cdot \sum_{k=0}^{\infty} \underbrace{\frac{\Gamma(\alpha + \beta + 1)}{\Gamma(\alpha + 1)\Gamma(\beta)} \frac{\Gamma(\alpha + 1 + k)\Gamma[\beta + (m - 1) - k]}{\Gamma[\alpha + 1 + \beta + (m - 1)]}}_{\sim BB(\alpha+1, \beta, m-1)} \frac{(m - 1)!}{(m - 1 - k)!k!} \\ &= \frac{m\alpha}{\alpha + \beta} = m\pi \end{aligned}$$

Beta-Binomial分布的期望和方差的另一种求法

- 为求 y 的方差，先求 $E(Y^2)$:

$$\begin{aligned} E(Y^2) &= \sum_{y=0}^{\infty} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + y)\Gamma(\beta + m - y)}{\Gamma(\alpha + \beta + m)} \\ &\quad \cdot \frac{m!}{(m - y)!y!} (y^2 - y + y) \\ &= \sum_{y=2}^{\infty} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + 2 + y - 2)\Gamma(\beta + m - 2 - (y - 2))}{\Gamma(\alpha + 2 + \beta + m - 2)} \\ &\quad \cdot \frac{m(m - 1)(m - 2)!}{[m - 2 - (y - 2)]!(y - 2)!} + E(y) \\ &= \frac{m(m - 1)\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} + \frac{m\alpha}{\alpha + \beta} \\ &= \frac{m\alpha[(m - 1)(\alpha + 1) + \alpha + \beta + 1]}{(\alpha + \beta)(\alpha + \beta + 1)} = \frac{m\alpha(m\alpha + m + \beta)}{(\alpha + \beta)(\alpha + \beta + 1)} \end{aligned}$$

Beta-Binomial分布的期望和方差的另一种求法

● 则有：

$$\begin{aligned} \text{Var}(Y) &= \frac{m\alpha(m\alpha + m + \beta)}{(\alpha + \beta)(\alpha + \beta + 1)} - \frac{m^2\alpha^2}{(\alpha + \beta)^2} \\ &= \frac{m\alpha(m\alpha + m + \beta)(\alpha + \beta) - m^2\alpha^2(\alpha + \beta + 1)}{(\alpha + \beta)^2(\alpha + \beta + 1)} \\ &= \frac{m\alpha\beta(\alpha + m + \beta)}{(\alpha + \beta)^2(\alpha + \beta + 1)} \\ &= m\left(\frac{\alpha}{\alpha + \beta}\right)\left(\frac{\beta}{\alpha + \beta}\right)\left(\frac{\alpha + \beta + m}{\alpha + \beta + 1}\right) \\ &= m\pi(1 - \pi)\left(\frac{\alpha + \beta + m}{\alpha + \beta + 1}\right) \\ &> m\pi(1 - \pi) \end{aligned}$$

Binomial分布过离散的直观解释

由前面的结果可知:

$$\begin{aligned}\text{Var}(Y) &= m\left(\frac{\alpha}{\alpha+\beta}\right)\left(\frac{\beta}{\alpha+\beta}\right)\left(\frac{\alpha+\beta+1+m-1}{\alpha+\beta+1}\right) \\ &= m\left(\frac{\alpha}{\alpha+\beta}\right)\left(\frac{\beta}{\alpha+\beta}\right)\left(1+\frac{m-1}{\alpha+\beta+1}\right)\end{aligned}$$

令 $\frac{1}{\alpha+\beta+1} = \rho, \frac{\alpha}{\alpha+\beta} = \pi$ 则有:

$$\begin{aligned}\text{Var}(Y) &= m\left(\frac{\alpha}{\alpha+\beta}\right)\left(\frac{\beta}{\alpha+\beta}\right)[1+\rho(m-1)] \\ &= m\pi(1-\pi)[1+\rho(m-1)]\end{aligned}$$

Binomial分布过离散的直观解释

设 $Y = \sum_{j=1}^m Y_j$, 其中 $Y_j \sim B(1, \pi)$, 且 $\text{Corr}(y_i, y_j) = \begin{cases} \rho & i \neq j \\ 1 & i = j \end{cases}$, 则有:

$$E(Y) = E\left(\sum_{j=1}^m Y_j\right) = m\pi(1 - \pi)$$

$$\begin{aligned} \text{Var}(Y) &= \text{Var}\left(\sum_{j=1}^m Y_j\right) = \sum_{j=1}^m \text{Var}(Y_j) + 2 \sum_{i=1}^m \sum_{j \neq i}^m \text{Cov}(y_i, y_j) \\ &= m\pi(1 - \pi) + m(m - 1)\rho\pi(1 - \pi) \\ &= m\pi(1 - \pi)(1 + (m - 1)\rho) \end{aligned}$$

∴ 当Y服从Beta-二项分布时, 可以看成是m个具有相关性的伯努利随机变量之和, 其相关系数为 ρ .

广义线性回归

- 现在考虑广义线性回归的情况：
- 设 $p_i \sim \text{Beta}(a_i, b_i)$, $Y_i|p_i \sim \text{Binomial}(m, p_i)$, 这里假设 m 为常量。
- 则 p_i 的先验分布均值为：

$$\mu_{p_i} = \frac{\alpha_i}{\alpha_i + \beta_i}$$

- 考虑 *logit* 链接函数：

$$\log\left(\frac{\mu_{p_i}}{1 - \mu_{p_i}}\right) = a_0 + a_1 x_i$$

广义线性回归

- 设 $\rho = \frac{1}{\alpha_i + \beta_i + 1}$, 则有: $\alpha_i + \beta_i = \frac{1}{\rho} - 1 = \frac{1 - \rho}{\rho}$, 且

$$\text{Var}(Y_i) = m\pi_i(1 - \pi)(1 + (m - 1)\rho)$$

- 从而:

$$\begin{aligned}\alpha_i &= mu_{p_i}(\alpha_i + \beta_i) = \frac{\rho}{1 - \rho} mu_{p_i} = \frac{1 - \rho}{\rho} \frac{e^{\alpha_0 + \alpha_1 x_i}}{1 + e^{\alpha_0 + \alpha_1 x_i}} \\ \beta_i &= \frac{1 - \rho}{\rho} - \alpha_i = \frac{1 - \rho}{\rho} (1 - mu_{p_i}) = \frac{1 - \rho}{\rho} \frac{1}{1 + e^{\alpha_0 + \alpha_1 x_i}}\end{aligned}$$

- 由此可以得到 $p_i \sim \text{Beta}(\frac{1 - \rho}{\rho} \frac{e^{\alpha_0 + \alpha_1 x_i}}{1 + e^{\alpha_0 + \alpha_1 x_i}}, \frac{1 - \rho}{\rho} \frac{1}{1 + e^{\alpha_0 + \alpha_1 x_i}})$
进而由 p_i 得到: $Y_i | p_i \sim \text{Binomial}(m, p_i)$

关于Beta-Binomial分布的拟似然

- 对于二项分布的广义线性模型

① $\{Y_1, \dots, Y_n\}$ 是一列独立的随机变量，服从二项分布 $B(m_i, p_i)$ ，这里我们设置为 $m_i = m$ 。

② $\mu_i = E(Y_i) = mp_i$ 是一列协变量 x_i 和参数向量 β 的组合，通过链接函数 $\eta_i = \log\left(\frac{\mu_i}{m - \mu_i}\right) = x_i^T \beta$ 。

✓ ③ 一般情况下 $\text{var}(Y_i) = V_i(\mu_i) = mp_i(1 - p_i)$ ，如果是过离散的情况 $\text{var}(Y_i) = \sigma^2 V_i(\mu_i) = \sigma^2 mp_i(1 - p_i)$ 。

关于二项分布的拟似然

- 对于随机变量 Y 拟似然得分函数 (QS) 为:

$$U = U(\mu; Y) = \frac{Y - \mu}{\sigma^2 V(\mu)}$$

- 故对于过离散的 Y 服从二项分布 $B(m, p)$,
 $V(\mu) = \frac{\mu(m - \mu)}{m}$, 它的拟似然得分函数 (QS) 可以表示为:

$$U = \frac{m(Y - \mu)}{\sigma^2 \mu(m - \mu)}$$

关于二项分布的拟似然

- 由一般的拟似然函数 (QL),

$$Q = Q(\mu; Y) = \int_Y^{\mu} \frac{Y - t}{\sigma^2 V(t)} dt$$

- 可以得到过离散情况下的二项分布的拟似然函数 (QL) 为

$$\begin{aligned} Q &= Q(\mu, Y) = \int_Y^{\mu} \frac{m(Y - t)}{\sigma^2 t(m - t)} dt \\ &= \frac{Y}{\sigma^2} \log\left(\frac{\mu}{Y}\right) - \frac{m - Y}{\sigma^2} \log\left(\frac{m - Y}{m - \mu}\right) \end{aligned}$$

关于二项分布的拟似然

- 故由一般的拟似然残差 (QD) 定义为:

$$\begin{aligned} D(y; \mu) &= 2\sigma^2 [Q(y; y) - Q(\mu; y)] \\ &= -2\sigma^2 Q(\mu; y) \end{aligned}$$

- 过离散的二项分布的拟似然残差 (QD) 为:

$$D(y; \mu) = -2[y \log(\frac{\mu}{y}) + (y - m) \log(\frac{m - y}{m - \mu})]$$

关于二项分布的拟似然

- 故对于独立变量 $Y = (Y_1, \dots, Y_n)^T$, $Y_i \sim B(m, p_i)$, 均值向量 $\mu = (\mu_1, \dots, \mu_n)^T = (mp_1, \dots, mp_n)^T$, Y 关于 μ 的QL为

$$Q(\mu; Y) = \sum_{i=1}^n \left(\frac{Y_i}{\sigma^2} \log\left(\frac{\mu_i}{Y_i}\right) + \frac{Y_i - m}{\sigma^2} \log\left(\frac{m - Y_i}{m - \mu_i}\right) \right)$$

- QS为 $U(\mu, Y) = \sum_{i=1}^n \frac{m(Y_i - \mu_i)}{\sigma^2 \mu_i (m - \mu_i)}$
- QD为 $D(\mu; Y) = -2 \sum_{i=1}^n \left[y \log\left(\frac{\mu}{y}\right) + (y - m) \log\left(\frac{m - y}{m - \mu}\right) \right]$

Poisson overdispersion

- One example of Poisson overdispersion is when we observe a random sum (in terms of how many terms we include) of iid RVs.
- **The cluster effects:** Let $\{Z_j : j = 1, \dots\}$ be a set of iid RVs where each Z_j has mean $E(Z)$ and variance $\text{var}(Z)$.
- **The clustering mechanism:** Suppose that N has a $\text{Poisson}(\lambda)$ distribution which is independent of each $Z_j, j = 1, \dots$.
- **What we observe:** Let $Y = \sum_{j=1}^N Z_j$.
- What are the first two moments Y ?

The expectation of Y



The variance of Y



A parametric model for Poisson overdispersion

(Breslow, 1984)

- Consider the following hierarchical model for Y :

$$Y|\lambda \sim \text{Poisson}(\lambda);$$

$$\lambda \sim \text{Gamma}(\alpha, \beta).$$

- What is the distribution of Y ?
- We have that

$$f(y|\lambda) = \frac{e^{-\lambda} \lambda^y}{y!}$$

and

$$f(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}$$

The joint distribution of Y and λ

$$\begin{aligned} f(y, \lambda) &= f(y|\lambda)f(\lambda) \\ &= \frac{e^{-\lambda}\lambda^y}{y!} \times \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \\ &= \frac{\beta^\alpha \lambda^{y+\alpha-1} e^{-(\beta+1)\lambda}}{\Gamma(\alpha)y!} \end{aligned}$$

The distribution of Y

$$\begin{aligned}f(y) &= \int_0^\infty f(y, \lambda) d\lambda \\&= \int_0^\infty \frac{\beta^\alpha \lambda^{y+\alpha-1} e^{-(\beta+1)\lambda}}{\Gamma(\alpha)y!} d\lambda \\&= \int_0^\infty \frac{(\beta+1)^{y+\alpha} \lambda^{y+\alpha-1} e^{-(\beta+1)\lambda}}{\Gamma(y+\alpha)} d\lambda \times \frac{\beta^\alpha \Gamma(y+\alpha)}{\Gamma(\alpha)(\beta+1)^{y+\alpha} y!} \\&= \frac{\Gamma(y+\alpha)}{\Gamma(\alpha)\Gamma(y+1)} \times \left(\frac{\beta}{\beta+1}\right)^\alpha \left(\frac{1}{\beta+1}\right)^y\end{aligned}$$

This is a Negative-Binomial pmf with parameters $r = \alpha$ and $p = \frac{\beta}{\beta+1}$ (see the midterm exam).

The moments of Y



A Negative Binomial example

- We consider a dataset titled 'Soldering of Components on Printed-Circuit Boards', taken from 'Statistical Models in S (1992), by John M. Chambers and Trevor J. Hastie eds, Wadsworth and Brooks/Cole, Pacific Grove, CA'
- The data has 720 rows and 6 columns, representing a balanced subset of a designed experiment varying 5 factors on the soldering of components on printed-circuit boards.
- We are interested in modeling the number of visible solder skips in terms of these 5 factors.

The factor variables

- 1. **Opening**: a factor with levels L, M, and S, indicating the amount of clearance around the mounting pad.
- 2. **Solder**: a factor with levels Thick and Thin, giving the thickness of the solder used.
- 3. **Mask**: a factor with levels A1.5, A3, B3, and B6 indicating the type and thickness of mask used.
- 4. **PadType**: a factor with levels D4, D6, D7, L4, L6,L7, L8, L9, W4, and W9 giving the size and geometry of the mounting pad.
- 5. **Panel**: values 1, 2, and 3, indicating the panel on a board being tested.

Choosing a model

- Since we are modeling **counts**, a Poisson GLM seems reasonable.
- Exercise:
 - Produce summaries of this data.
 - Find a **suitable** model for the number of visible solder skips in terms of the five factors.
- We will consider a simple model inspired by what Chambers and Hastie did.

A Poisson GLM

```
glm(formula = skips ~ Opening + Solder + Mask + PadType + factor(Panel), family = poisson, data = solder)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.6615	-1.0868	-0.4407	0.6115	3.9429

Coefficients:

	Estimate	Std Error	z value	$Pr(> z)$	
(Intercept)	-1.21987	0.09520	-12.813	$< 2e-16$	***
OpeningM	0.25851	0.06656	3.884	0.000103	***
OpeningS	-0.0073541	0.0016779	-4.383	0.000625	***
SolderThin	0.0190104	0.0016930	11.229	2.18e-08	***
MaskA3	0.42819	0.07547	5.674	1.40e-08	***
MaskB3	1.20225	0.06697	17.953	$< 2e-16$	***
MaskB6	1.86648	0.06310	29.580	$< 2e-16$	***
PadTypeD6	-0.36865	0.07138	-5.164	2.41e-07	***
PadTypeD7	-0.09844	0.06620	-1.487	0.137001	
PadTypeL4	0.26236	0.06071	4.321	1.55e-05	***
PadTypeL6	-0.66845	0.07841	-8.525	$< 2e-16$	***
PadTypeL7	-0.49021	0.07406	-6.619	3.61e-11	***
PadTypeL8	-0.27115	0.06939	-3.907	9.33e-05	***
PadTypeL9	-0.63645	0.07759	-8.203	2.35e-16	***
PadTypeW4	-0.11000	0.06640	-1.657	0.097590	
PadTypeW9	-1.43759	0.10419	-13.798	$< 2e-16$	***
factor(Panel)2	0.33352	0.04207	7.929	2.21e-15	***
factor(Panel)3	0.25440	0.04278	5.947	2.74e-09	***

3.1 The Analysis of Deviance

Analysis of Deviance Table

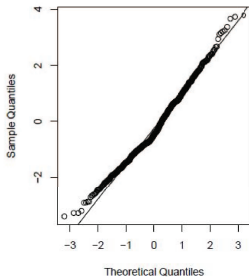
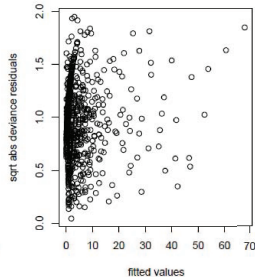
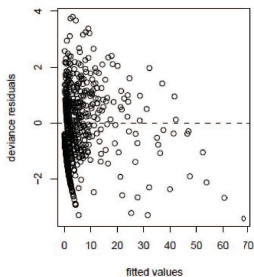
Model: poisson, link: log

Response: skips

Terms added sequentially (first to last)

	Df Deviance	Resid.	Df Resid.	Dev
NULL			719	6855.7
Opening	2	2524.6	717	4331.1
Solder	1	937.0	716	3394.2
Mask	3	1653.1	713	1741.1
PadType	9	542.5	704	1198.6
factor(Panel)	2	68.1	702	1130.5

Diagnostic plots



Estimating the dispersion parameter

- From the Pearson χ^2 statistic we estimate σ^2 by $\hat{\sigma}^2 = 1.49$
- Some evidence of overdispersion. Adjusting for overdispersion the coefficient summary becomes:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.21987	0.11623	-10.49493	0.00000
OpeningM	0.25851	0.08127	3.18102	0.00147
OpeningS	1.89349	0.06548	28.91712	0.00000
SolderThin	1.09973	0.04717	23.31405	0.00000
MaskA3	0.42819	0.09214	4.64718	0.00000
MaskB3	1.20225	0.08176	14.70444	0.00000
MaskB6	1.86648	0.07704	24.22751	0.00000
PadTypeD6	-0.36865	0.08715	-4.23001	0.00002
PadTypeD7	-0.09844	0.08082	-1.21798	0.22323
PadTypeL4	0.26236	0.07412	3.53954	0.00040
PadTypeL6	-0.66845	0.09573	-6.98242	0.00000
PadTypeL7	-0.49021	0.09042	-5.42147	0.00000
PadTypeL8	-0.27115	0.08472	-3.20040	0.00137
PadTypeL9	-0.63645	0.09473	-6.71845	0.00000
PadTypeW4	-0.11000	0.08107	-1.35689	0.17482
PadTypeW9	-1.43759	0.12721	-11.30130	0.00000
factor(Panel)2	0.33352	0.05136	6.49406	0.00000
factor(Panel)3	0.25440	0.05223	4.87052	0.00000

Fitting a negative binomial model to the data

- Suppose $\{Y_i : i = 1, \dots, n\}$ are a set of independent RVs with $Y_i \sim \text{NegBin}(r, p_i)$ for each i .
- From the midterm we saw that when r is fixed, we have an exponential family.
- Then in R we can fit the model using glm:
`model2 <- glm(skips ~ Opening + Solder + Mask + PadType + factor(Panel), data=solder, family=negative.binomial(theta=1))`
- Here we set $r=1$ (R calls the parameters θ not r)
- R assumes a **log link**, i.e., we model $\log \mu_i$, where $\mu_i = E(Y_i)$.

The model summary

glm(formula = skips ~ Opening + Solder + Mask + PadType + factor(Panel), family = negative.binomial($\theta=1$), data = solder) Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-1.9726	-0.7405	-0.2167	0.3115	2.4259
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.838167	0.167161	-10.996	<2e-16	***
OpeningM	0.429035	0.096954	4.425	1.12e-05	***
OpeningS	2.096484	0.089994	23.296	<2e-16	***
SolderThin	1.303673	0.073479	17.742	<2e-16	***
MaskA3	0.532464	0.114747	4.640	4.15e-06	***
MaskB3	1.313192	0.109464	11.997	<2e-16	***
MaskB6	2.120601	0.106596	19.894	<2e-16	***
PadTypeD6	-0.449175	0.159908	-2.809	0.005108	**
PadTypeD7	0.042471	0.154213	0.275	0.783085	
PadTypeL4	0.583751	0.149748	3.898	0.000106	***
PadTypeL6	-0.399470	0.159248	-2.508	0.012349	*
PadTypeL7	-0.280122	0.157744	-1.776	0.076199	
PadTypeL8	-0.012017	0.154759	-0.078	0.938128	
PadTypeL9	-0.480472	0.160335	-2.997	0.002826	**
PadTypeW4	0.008515	0.154551	0.055	0.956078	
PadTypeW9	-1.403285	0.177458	-7.908	1.02e-14	***
factor(Panel)2	0.324748	0.088425	3.673	0.000258	***
factor(Panel)3	0.314383	0.088483	3.553	0.000406	***

(Dispersion parameter for Negative Binomial(1) family taken to be 0.5579386)

The analysis of deviance

Analysis of Deviance Table

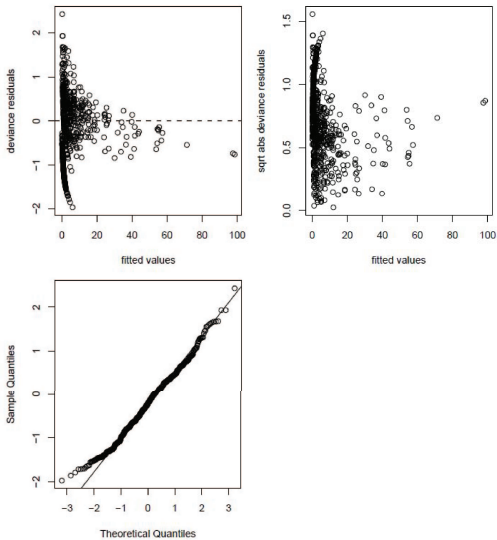
Model: Negative Binomial(1), link: log

Response: skips

Terms added sequentially (first to last)

	Df	Deviance	Resid.	Df	Resid.	Dev
NULL				719		1423.13
Opening	2	435.34		717		987.79
Solder	1	173.44		716		814.35
Mask	3	292.97		713		521.38
PadType	9	91.37		704		430.01
factor(Panel)	2	9.34		702		420.67

Diagnostic plots



Remarks on this model

- Coefficient summaries as compared with the Poisson model:
- Dispersion parameter:
- Residual deviance for this model compared to the Poisson model:

7.1 Estimating the $r(\theta)$ parameter

- To estimate the r parameter we use the following R code (which also requires the **MASS** library):
model3 <- glm.nb(skips ~ Opening + Solder + Mask + PadType + factor(Panel), data=solder)
summary(model3, cor=F)
- This code runs very slowly!

The model summary

```
glm.nb(formula = skips ~ Opening + Solder + Mask + PadType +  
factor(Panel), family = negative.binomial( $\theta=1$ ), data = solder)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-2.8944	-0.9552	-0.3921	0.5304	3.2348
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.49846	0.12472	-12.014	<2e-16	***
OpeningM	0.30623	0.07840	3.906	9.39e-05	***
OpeningS	2.00244	0.06686	29.950	<2e-16	***
SolderThin	1.21356	0.05236	23.179	<2e-16	***
MaskA3	0.45408	0.09112	4.983	6.25e-07	***
MaskB3	1.22575	0.08345	14.689	<2e-16	***
MaskB6	2.00400	0.07945	25.224	<2e-16	***
PadTypeD6	-0.43954	0.10645	-4.129	3.64e-05	***
PadTypeD7	-0.05931	0.10023	-0.592	0.5540	
PadTypeL4	0.39075	0.09470	4.126	3.69e-05	***

The Analysis of Deviance

Analysis of Deviance Table

Model: Negative Binomial(10.5345), link: log

Response: skips

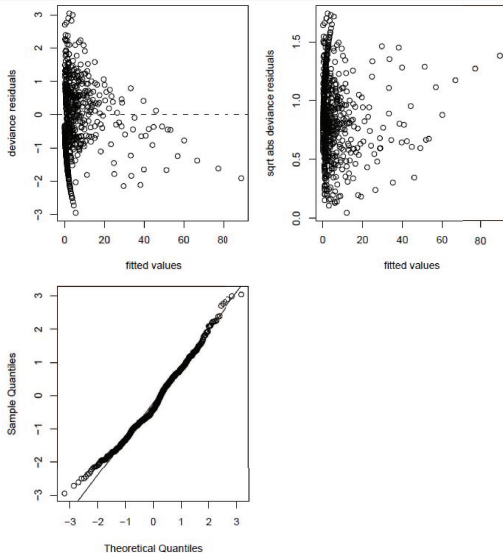
Terms added sequentially (first to last)

	Df	Deviance	Resid.	Df	Resid.	Dev	P(> Chi)
NULL				719		4488.1	
Opening	2		1699.0	717		2789.1	0.0
Solder	1		581.0	716		2208.1	2.278e-128
Mask	3		1014.2	713		1193.9	1.493e-219
PadType	9		313.6	704		880.3	3.316e-62
factor(Panel)	2		37.7	702		842.6	6.526e-09

Warning message:

tests made without re-estimating theta in: anova.negbin(model3)

Diagnostic plots



Conclusions

- Coefficient summaries as compared with the Poisson model:
- Dispersion parameter:
- Residual deviance for this model compared to the Poisson model:
- Overall comments:

The Gamma distribution

- Suppose a R.V. Y has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$.

We say $Y \sim \text{Ga}(\alpha, \beta)$.

- The pdf of Y at a value y is given by

$$f_Y(y) = \frac{y^{\alpha-1} e^{-y/\beta}}{\Gamma(\alpha) \beta^\alpha}, y \geq 0$$

- The **moment generating function** of Y is

$$M_Y(t) = (1 - \beta t)^{-\alpha}$$

- Thus, the **mean** is

$$E(Y) = \alpha\beta,$$

and the **variance** is

$$\text{var}(Y) = \alpha\beta^2.$$

Special cases of the Gamma distribution

1. $\alpha = 1$: Exponential distribution.
2. $\alpha = p/2$ and $\beta = 2$: Chi-squared(p) distribution.
3. $\alpha = 1$, and let $Z = Y^{1/\gamma}$ for some $\gamma > 0$: Then Z has a Weibull distribution with parameters γ and β .

Why do we need the Gamma distribution?

Re-parameterizing the Gamma distribution

- We want to redefine our parameters so that

$$E(Y) = \alpha\beta = \mu,$$

and

$$\text{var}(Y) = \alpha\beta^2 = \frac{\mu^2}{\nu}.$$

Solving, we have that $\alpha = \nu$ and $\beta = \mu/\nu$, and hence

$$Y \sim \text{Ga}(\nu, \mu/\nu).$$

- Then can show that the pdf of Y in terms of this parameterization is

$$f_Y(y) = \frac{(y\nu)^\nu \exp(-\frac{y\nu}{\mu})}{y\Gamma(\nu)\mu^\nu}, y \geq 0$$

Do we have an exponential family?

• Let

$$\theta = -\frac{1}{\mu}$$

$$a(\phi) = \frac{1}{\nu}$$

$$b(\theta) = -\log(-\theta)$$

$$c(y, \phi) = \nu \log(\nu y) - \log(y) - \log(\Gamma(\nu))$$

$$f_Y(y) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right)$$

Checking the moments

- The expectation of Y is

$$E(Y) = b'(\theta) = -\frac{1}{\theta} = \mu.$$

Since

$$b''(\theta) = \frac{1}{\theta^2} = \mu^2,$$

we have that $\text{var}(Y) = b''(\theta)a(\phi) = \mu^2/\nu$.

- We can see the mean-variance relationship from the fact that

$$V(\mu) = \mu^2.$$

GLMs for Gamma data

- Suppose $\{Y_1, \dots, Y_n\}$ is a set of independent $Ga(\mu_i, \mu_i/\nu)$ RVs.
- The canonical link for this family is

$$\eta_i = g(\lambda_i) = \frac{1}{\mu_i}$$

- Regardless of the link function $g()$ chosen, we fit the model

$$\eta_i = \mathbf{x}_i^T \beta, \quad i = 1, \dots, n.$$

- We estimate the parameters β using IWLS: as in the Poisson case, we need to be careful about the choice of starting values.

Choices of link function

1. The canonical link:

$$\eta_i = \frac{1}{\mu_i}$$

We must restrict the set of $\hat{\beta}$ values so that $\hat{\mu} > 0$, for all i .

2. The log link:

$$\eta_i = \log \mu_i.$$

No restriction on the set of $\hat{\beta}$ values in this case.

3. The identity link:

$$\eta_i = \mu_i.$$

Again we must restrict the set of $\hat{\beta}$ values so that $\hat{\mu}_i > 0$, for all i .

The deviance statistic

- Since $\theta_i = -1/\mu_i$ we have that

$$\tilde{\theta}_i = -\frac{1}{y_i}$$

, and

$$\hat{\theta}_i = -\frac{1}{\hat{\mu}_i},$$

and using $b(\theta_i) = -\log(-\theta_i) = \log(\mu_i)$, we get

$$b(\tilde{\theta}_i) = \log(y_i),$$

and

$$b(\hat{\theta}_i) = \log(\hat{\mu}_i).$$

- Thus

$$D(y, \mu) = 2 \sum_{i=1}^n \left[\left(\frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} \right) - \log \left(\frac{y_i}{\hat{\mu}_i} \right) \right]$$

The Pearson statistic, and an estimate of ν

- The Pearson χ^2 statistic is

$$\begin{aligned}\chi^2 &= \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)} \\ &= \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{(\hat{\mu}_i)^2}\end{aligned}$$

- We can estimate $\sigma^2 = \nu^{-1}$ using the standard moment estimator,

$$\hat{\sigma}^2 = \frac{\chi^2}{n - p}.$$

Hence

$$\hat{\nu} = \frac{n - p}{\chi^2}.$$

- Alternatively, we could use the MLE of ν - this estimator is biased for small n (it is possible to correct the bias of the MLE).

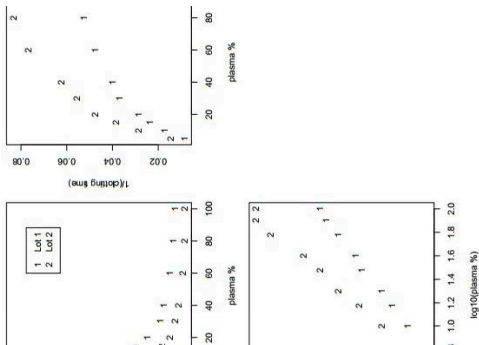
Example: clotting times of blood

“Hurn, et al. (1945) published data on the clotting time of blood, giving clotting time in seconds for normal plasma diluted to nine different percentage concentrations with prothrombin-free plasma; clotting was induced by two lots of thromboplastin.”

(Taken from section 8.4.2 of McCullagh and Nelder)

- Scientific aim: Model the clotting times in terms of the plasma concentration % and the lot number.

Summary plots



Fitting a Gamma GLM

- Let Y_{ij} be the clotting time for plasma concentration percentage, $\mu_i (i = 1, \dots, 9)$ and lot $j = 1, 2$.
- Assume $\{Y_{ij}\}$ and a set of independent $Ga(\mu_{ij}, \frac{\mu_{ij}}{\nu})$ RVs with

$$1/\mu_{ij} = \alpha + \beta \log_{10}(\mu_i) + \gamma_j + (\beta\gamma)_j \log_{10}(\mu_i)$$

- To fit the model we assume that

$$\gamma_1 = 0 \text{ and } (\beta\gamma)_1 = 0.$$

The model summary

Call :

```
glm(formula = clot.time ~ log.plasma * lot, family = Gamma)
```

DevianceResiduals :

Min	1 Q	Median	3 Q	Max
-0.055738	-0.035480	-0.008216	0.026073	0.086411

Coefficients :

	Estimate	Std. Error	tvalue	Pr(> t)
(Intercept)	-0.0165544	0.0008655	-19.127	1.97e - 11 * **
log.plasma	0.0353288	0.0008915	39.626	8.85e - 16 * **
lot2	-0.0073541	0.0016779	-4.383	0.000625 * **
log.plasma : lot2	0.0190104	0.0016930	11.229	2.18e - 08 * **

— — —

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.002129676)

Null deviance: 7.708667 on 17 degrees of freedom

Residual deviance: 0.029401 on 14 degrees of freedom

AIC: 63.195

Number of Fisher Scoring iterations: 3

The dispersion parameter

- In S-PLUS and R the Gamma GLM always estimates the dispersion parameter, $\sigma^2 = 1/\nu$. For the blood clotting example the estimate is

```
sigma2 <- summary(model)$dispersion  
sigma2
```

```
[1]0.002129676
```

- The standard errors, t-value and P-values are all based on this estimate on σ^2 (i.e., we adjust for the dispersion in the data).
- For this model:

The analysis of deviance

Analysis of Deviance Table

Model: Gamma, link: inverse

Response: clot.time

		<i>Df</i>	<i>Deviance</i>	<i>Resid. Df</i>	<i>Resid. Dev</i>
	<i>NULL</i>			17	7.7087
<i>Terms added sequentially (first to last)</i>	<i>log.plasma</i>	1	6.6904	16	1.0183
	<i>lot</i>	1	0.7178	15	0.3004
	<i>log.plasma : lot</i>	1	0.2710	14	0.0294

Comments on the analysis of deviance

- Since we estimate σ^2 , the deviance tests become:

```
1 - pchisq(6.6904 / sigma2, 1)
```

```
[1]0
```

```
1 - pchisq(0.7178 / sigma2, 1)
```

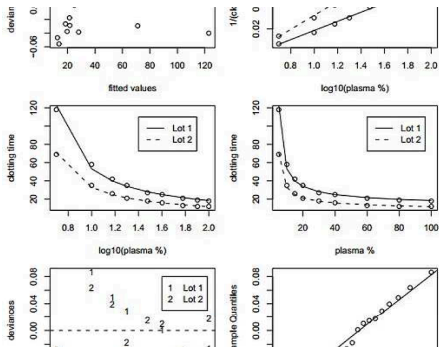
```
[1]0
```

```
1 - pchisq(0.2710 / sigma2, 1)
```

```
[1]0
```

- Conclusions:

Diagnostic plots



Comments on the diagnostic plots



Removing the first case, and refitting the Gamma GLM

*glm(formula = clot.time ~ log.plasma * lot, family = Gamma, subset = 2 : 18) Deviance Residuals:*

Min	1 Q	Median	3 Q	Max
-0.05574	-0.01754	0.00937	0.01436	0.06372

Coefficients:

	Estimate	Std. Error	tvalue	Pr(> t)
(Intercept)	-0.021772	0.001548	-14.061	3.06e - 09 * * *
log.plasma	0.038947	0.001186	32.851	6.77e - 14 * * *
lot2	-0.002136	0.001866	-1.145	0.273
log.plasma : lot2	0.015392	0.001578	9.753	2.40e - 07 * * *

**** 0.01 '** 0.05 '.' 0.1 ' ' 1*

(Dispersion parameter for Gamma family taken to be 0.001115850)

Null deviance: 4.571335 on 16 degrees of freedom

Residual deviance: 0.014508 on 13 degrees of freedom

AIC: 46.194

Number of Fisher Scoring iterations: 3

The analysis of deviance

- Analysis of Deviance Table

Model: Gamma, link: inverse

Response: clot.time

		Df	Deviance	Resid.Df	Resid.Dev
	NULL			16	4.5713
Terms added sequentially (first to last)	log.plasma	1	3.6365	15	0.9349
	lot	1	0.8145	14	0.1203
	log.plasma : lot	1	0.1058	13	0.0145

- Again we adjust for $\hat{\sigma}^2 = 0.00112$:

1 - pchisq(3.6365 / sigma2, 1)

[1]0

1 - pchisq(0.8145 / sigma2, 1)

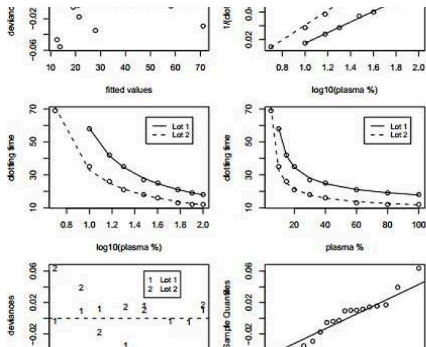
[1]0

1 - pchisq(0.1058 / sigma2, 1)

[1]0

- Conclusions:

Diagnostic plots



Summary of the model

