Chapter 18

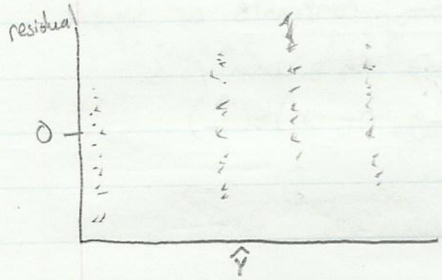
Residual $= e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - \bar{Y}_{i}$

Semistudentized residuals $e_{ij}^{*} = e_{ij} / \sqrt{MSE}$

Studentized Residual $r_{ij} = e_{ij} / s(e_{ij}) \ni s(e_{ij}) = \sqrt{\frac{MSE(n_i - 1)}{n_i}}$

Studentized Deleted Residual $t_{ij} = \frac{n_T - r - 1}{SSE(1 - \frac{1}{n_i}) - e_{ij}^2}$

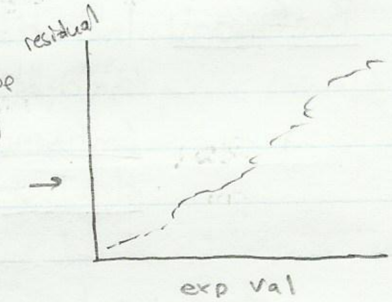
*Diagnostics from glimmix plots = all



← • no megaphone shape

← • balanced around 0

• fits prob plot line →



- We can also look at residuals vs factor levels

- Nonconstant error variance - if we see megaphone shape in residual vs $\hat{Y}$ plot
- NonIndependent error terms - if we see order or pattern in a residual sequence plot
- Outliers · Look at boxplots, residual vs. $\hat{Y}$ plot or do bonferroni test
- Are we missing an explanatory variable?
- Nonnormality of error terms look at residual normality plot

Formal tests for constant error variance Hartley or Brown-Forsythe Tests

Both test $H_0: \sigma_1^2 = \sigma_2^2 = \dots \sigma_r^2$

H_a: not all are equal

not available from SAS → Hartley: $H^* = \frac{\max(s_i^2)}{\min(s_i^2)}$ reject if $H^* > H(1-\alpha, r, df)$

Brown Forsythe: $F_{BF}^{*} = \frac{MStreatment}{MSE}$ reject if F_{BF}^{*} is large

SAS
```
proc glm; class x;
model y=x;
means x / HOVTEST=BF;
run;
```

Weighted Least Squares

$w_{ij} = \frac{1}{s_i^2}$ ∋ $s_i^2 = \frac{\Sigma (Y_{ij} - \bar{Y}_{i\cdot})^2}{n_i - 1}$

$H_0: \mu_1 = \mu_2 = \dots \mu_n$

H_a: Not all μ_i are equal

Full model: $Y_{ij} = \mu_1 X_{ij1} + \mu_2 X_{ij2} + \dots + \mu_n X_{ijn} + \epsilon_{ij}$

where $X_i = \begin{cases} 1 & \text{if case is from factor level } i \\ 0 \end{cases}$

$$F_w^* = \frac{SSE_w(R) - SSE_w(F)}{r-1} \div \frac{SSE_w(F)}{n_T - r}$$

for some reduced model R

SAS

```
*Step 1
proc glm data= __ ;
    class x1;
    model y= x1;
    output out= temp r=residual;
run;
*Step 2: data temp; set temp; absr = abs(residual); run
*Step 3:
proc glm data= temp;
    class x1;
    model absr = x1;
    output out= temp1 p= s_hat;
run;
*Step 4: data temp1; set temp1; w= 1/(s_hat**2); run;
*step 5:
proc glm data= temp1;
    class x1;
    weight w;
    model y = x1 /solution;
run;
```

Transformation of Resp. Variable Like before we will use Box Cox

SAS
```
proc transreg data=___;
    model boxcox(y) = identity(x1);
run;
```

Welch's Test - An adaptation of Student's t-test used for two samples w/ possibly unequal variance.

$H_0: \mu_1 = \mu_2 = \dots = \mu_n$

H_0: Not all μ_i are equal

Levene's Test - for equal variance accross samples.

SAS
```
proc glm data=___;
    class x1;
    model y=x1;
    means x1 /welch hovtest; * gives Levene and Welch
run;
```

Rank F test * Y_{ij} observations must be ranked in ascending order (Rank = R_{ij})

$H_0: \mu_1 = \mu_2 = \dots \mu_n$

H_a: Not all μ_i are equal

$$F_R^* = \frac{MSTR}{MSE} = \frac{\sum n_i (\bar{R}_{i.} - \bar{R}_{..})^2}{r-1} \div \frac{\sum\sum (R_{ij} - \bar{R}_{i.})^2}{r_T - r}$$

Reject when $F_R^* > F(1-\alpha; r-1; n_T - r)$

Kruskal Wallis * Y_{ij} observations must be ranked in ascending order as above

$H_0: \mu_1 = \mu_2 = \dots \mu_n$

H_a: Not all μ_i are equal

$$\chi^2_{KW} = SSTR \div \frac{SSTO}{n_T - 1}$$

Reject when $\chi^2_{KW} > \chi^2(1-\alpha, r-1)$

<u>Poisson Regression Model</u> Poisson w/ log-link for count data

[SAS]
```
proc genmod;
    class x1;
    model y = x1 / dist = poisson link = log type3;
run;
```

<u>Multiple Pairwise Testing</u> Bonferroni pairwise comparisons to combine levels

$g = \frac{r(r-1)}{2}$

$B = z(1-\alpha/2g)$

$= \bar{R}_{i\cdot} - \bar{R}_{i'\cdot} \pm B\left[\frac{n_T(n_T+1)}{12}\left(\frac{1}{n_i}+\frac{1}{n_{i'}}\right)\right]^{1/2}$

[SAS]
```
proc glimmix data=__;
    class x1;
    model y=x1;
    lsmestimate x1 "M1-M2" 1 -1 0 .. 0 / adjust=bonf alpha=.01;
run;
```

[SAS]
```
/*Kruskal Wallace*/
proc nparlway data=____ WILCOXON;
    class x1
    EXACT WILCOXON /MC
    VAR Y;
run;
```