$\mu = EY$ theoretical mean $\Leftarrow$ property of model

$\bar{y} = \frac{1}{n}\sum y_i$ sample mean

If model density of Y is $f(y)$

$\mu = \int y f(y)\,dy$ by def$^{=}$

HW Qu 1 $\quad y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$

$\varepsilon_{ij} \sim (0, \sigma^2)$

$\Rightarrow \underbrace{E y_{ij}}_{\mu_{ij}} = \mu + \tau_i + \beta_j \qquad \operatorname{var} y_{ij} = \sigma^2$

theoretical

lm produces estimates $\hat{\mu}, \hat{\tau}_1, \dots, \hat{\tau}_5,$
aov $\hat{\beta}_1, \dots, \hat{\beta}_3$

actually 2 fewer

l.s. $\min\limits_{\mu, \tau_1, \dots \tau_5, \beta_1 \dots \beta_3} \sum_i \sum_j (y_{ij} - \mu - \tau_i - \beta_j)^2$

$5+3+1$ eq$^{=}$s but 2 lin. dep.

most 'natural' is to have $\hat{\mu} = \bar{y}_{..}$

$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$$

$$\hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..}$$

LS under constraints $\Sigma \tau_i = 0$, $\Sigma \beta_j = 0$

(R default constraint is something else

```
>options(contrasts=c("contr.sum",
                     "contr.poly"))
```

```
tmt <- factor(c(1,2,3,4,5, 1,2,3,4,5, 1,2,3,4,5))
```

X built by lm incorporates the constraints

$\mu \quad \tau_i \quad \beta_j$ are not separately estimable

BUT $\tau_i - \tau_{i'}$ is | $\sum_{i=1}^{5} l_i \tau_i$ is if $\Sigma l_i = 0$

$\mu + \tau_i$ is (?)

linear contrasts

Robust Regression

Recall $\sum_{i=1}^{n} \psi(y_i - \tilde{\mu}) = 0$ defines

an estimator $\tilde{\mu} = \tilde{\mu}(y)$ which estimates

$\mu = EY$ in a model in which $y_1, \dots, y_n$

are ind't obs$^{=}$ from the same dist$^{=}$

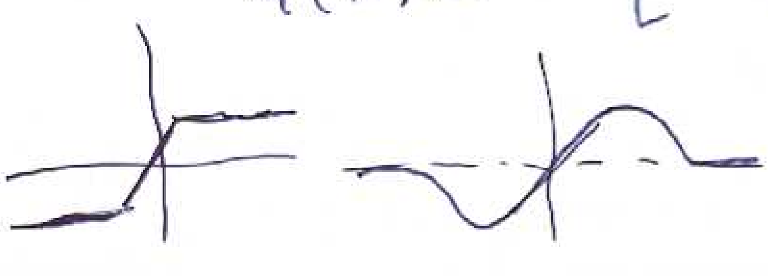
for some $f^{=}$ $\psi(\cdot)$

Huber's $f^{=}$ $\psi(x) = \begin{cases} -c & x < -c \\ x & -c \le x \le c \\ +c & x > c \end{cases}$

$c \simeq 1.5\tilde{\sigma}$ where $\tilde{\sigma}$ is an estimate of scale, usually MAD, IQR (or Huber 2)

Tukey's bisquare sometimes preferred See Fig 5.6

$$\psi(x) = x\left[1 - \left(\frac{x}{R}\right)^2\right]_+^2$$



Linear regression §6.5
Robust

$$(*) \quad \sum_{i=1}^{n} \psi\left(\frac{y_i - x_i^T \tilde{\beta}}{\sigma} \right) \cdot \underset{p\times 1}{x_i} = \underset{\sim}{0}_{p \times 1} \qquad \begin{array}{l} p \text{ eq's} \\ \& \ p \text{ unk.} \end{array}$$

$$E Y_i = \underset{\sim}{x}_i^T \underset{\sim}{\beta} \qquad (x_{i1}, \dots, x_{ip}) = x_i^T$$

<u>Motivation</u> $\quad y_i = x_i^T \beta + \sigma e_i \qquad e_i \sim f(\cdot)$

$$\ell(\beta) = \sum_{i=1}^{n} \log f\left(\frac{y_i - x_i^T \beta}{\sigma} \right) - n \log \sigma$$

(assume σ known)

$$\ell'(\beta) = -\sum g\left(\frac{y_i - x_i^T \beta}{\sigma} \right) \underset{\sim}{x}_i \Bigg|_{\tilde{\beta}} = 0$$

Sol^n by wt'd least squares $(*)$

$$(*) = 0 \iff \sum w_i \cdot \left(y_i - x_i^T \beta \right) \underset{\sim}{x}_i = \underset{\sim}{0}$$

$$w_i = \psi\left(\frac{y_i - x_i^T \beta}{\sigma} \right) \Big/ \left(\frac{y_i - x_i^T \beta}{\sigma} \right)$$

iterated to convergence

 $\Leftarrow$ hard to <u>prove</u> convergence

- in hills data

2 outliers

2 influential obs =

compare ~~lm~~ lm & rlm coefficients,
and predictions for these

- in book p. 162 several other models
fitted - such as weighted rlm

$$\frac{1}{dist^2}$$

- fitted $\frac{time}{dist}$ against $\frac{climb}{dist}$

$\uparrow$

1/speed

(intercept consistent
with 0)