Single Factor Experiments: Estimating the Parameters

- In the *means* model: $\widehat{\mu}_i = \bar{y}_{i\cdot}$.
- In the *effects* model, $\mu_i = \mu + \tau_i$, and we cannot estimate μ and τ_i separately without a constraint on the τ_i's.
- Natural constraint: $\sum \tau_i = 0$;
  - The intercept μ is then an overall mean.
- Most computer packages set one τ to zero.
  - The intercept μ is then the mean for that *baseline* level.

```
> summary(lm(EtchRate ~ factor(Power), data = etchRate))

Call:
lm(formula = EtchRate ~ factor(Power), data = etchRate)

Residuals:
   Min     1Q Median     3Q    Max
 -25.4  -13.0    2.8   13.2   25.6

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       551.200      8.169  67.471  < 2e-16 ***
factor(Power)180   36.200     11.553   3.133  0.00642 **
factor(Power)200   74.200     11.553   6.422 8.44e-06 ***
factor(Power)220  155.800     11.553  13.485 3.73e-10 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 18.27 on 16 degrees of freedom
Multiple R-Squared: 0.9261,     Adjusted R-squared: 0.9122
F-statistic:  66.8 on 3 and 16 DF,  p-value: 2.883e-09
```

```
> summary(lm(EtchRate ~ 0 + factor(Power), etchRate))

Call:
lm(formula = EtchRate ~ 0 + factor(Power), data = etchRate)

Residuals:
   Min      1Q  Median      3Q     Max
 -25.4   -13.0     2.8    13.2    25.6

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
factor(Power)160  551.200      8.169   67.47   <2e-16 ***
factor(Power)180  587.400      8.169   71.90   <2e-16 ***
factor(Power)200  625.400      8.169   76.55   <2e-16 ***
factor(Power)220  707.000      8.169   86.54   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 18.27 on 16 degrees of freedom
Multiple R-Squared: 0.9993,     Adjusted R-squared: 0.9991
F-statistic:  5768 on 4 and 16 DF,  p-value: < 2.2e-16
```

Interval Estimates

- One-at-a-time $100(1-\alpha)\%$ confidence intervals:
  - For a treatment mean μ_i:

$$\bar{y}_{i\cdot} \pm t_{\alpha/2,N-a}\sqrt{\frac{\mathsf{MS}_E}{n}}.$$

  - For a pairwise difference $\mu_i - \mu_j$:

$$\bar{y}_{i\cdot} - \bar{y}_{j\cdot} \pm t_{\alpha/2,N-a}\sqrt{\frac{2\mathsf{MS}_E}{n}}.$$

A problem: multiplicity

- Each interval is an assertion: "$\mu_i - \mu_j$ lies between two values".

- The probability that the assertion is wrong is α–often 5%.

- If we make many such assertions, we should expect around 5% of them to be wrong.

- If we make *all pairwise comparisons* among a treatments, we make $a(a-1)/2$ assertions. E.g. for $a = 7$ that's 21 assertions, so we expect roughly one to be wrong.

A Solution: Simultaneous Confidence Intervals

- The probability that *all* of a collection of confidence intervals are correct is the *experiment-wise* error rate.

- The probability that *one* confidence interval is correct is the *comparison-wise* error rate. Sometimes we use confidence intervals with a given comparison-wise error rate, like those on an earlier slide.

- Sometimes, e.g. when we need to *rank* the treatments, we need to control the experiment-wise error rate.

- Bonferroni method, Scheffé method, Dunnett method.

Balance; Unbalanced Designs

- A design with n runs for each level of the factor is *balanced*.

- Sometimes we may have different sample sizes: n_i runs for level i. Formulas for F_0 and CIs change to reflect lack of balance.

  - By chance: e.g. one or more measurements has to be excluded from the analysis because of incorrect procedures.

  - By design: if we are interested in making comparisons between one distinguished level, the *control* level, and each of the others, we may use more runs at the control level, to gain precision.

Model Adequacy Checking

- Using the F-distribution to assess F_0 and using the t-distribution to set up CIs both depend on the data being normally distributed with constant variance.

- Checks are based on *residuals*

$$e_{i,j} = y_{i,j} - \widehat{y}_{i,j}$$

  where in the single-factor model $\widehat{y}_{i,j} = \widehat{\mu}_i = \bar{y}_{i\cdot}$.

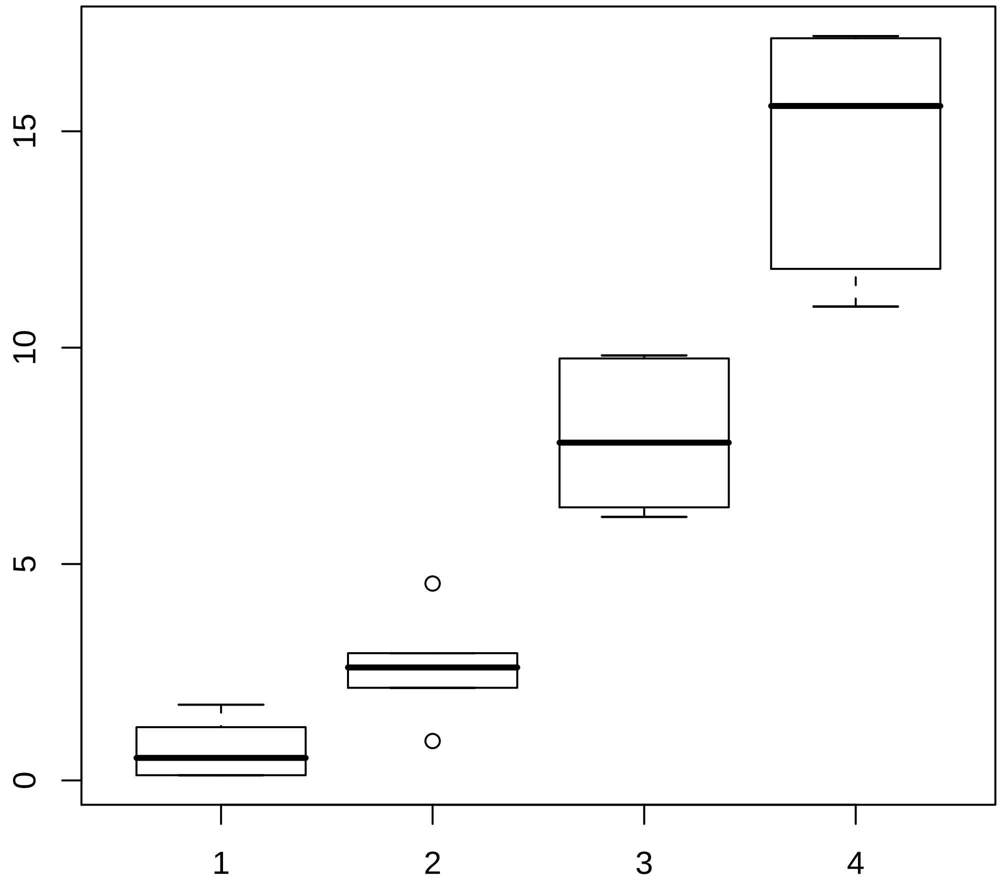
- Probability plot (quantile-quantile plot) of *pooled* residuals: may reveal outliers, other departures from normality.

- Time plot (residuals in time order): may reveal *correlation*.

- Residuals versus fitted values: may reveal non-constant variance; e.g. variance changes systematically with mean.

  – Can often correct by *transformation*: most commonly logarithms, but also square roots, other powers (Box-Cox).

Peak discharge rates



Peak discharge rates (sqrt scale)