Lecture 20 Wednesday March 1

PLAN

- Bonferroni method of multiple comparisons
- Discuss Project 1 Q#3B conclusions from residual plots (done verbally in lecture - no written notes)
- Illustrate R code (in Lec1-extra.r) for adding superimposed CIs to aligned dot plots for Q#4 in Project 1.

Preliminary note on the Bonferroni method

Bonferroni's inequality gives a formula for an upper bound on the probability of the union of events, where the events don't have to be mutually exclusive.

The probability of the union of $A_1, A_2, \ldots, A_k$ is the probability that "at least one of the events occurs."

FACT For any two events A and B, we know

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

by the addition rule. Thus, for A and B any events, we have

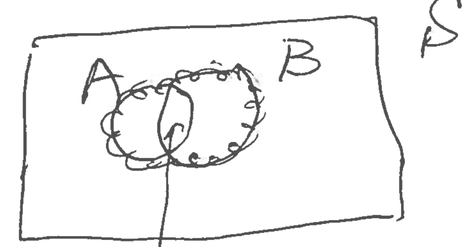
$$P(A \cup B) \leq P(A) + P(B)$$

In general:

$$\boxed{P\left(\cup_{i=1}^{k} A_i \right) \leq \Sigma_{i=1}^{k} P(A_i)}$$

Bonferroni inequality

Venn diagram



$A \cap B$

$$P(A) + P(B) = [P(A \cap B^c) + P(A \cap B)] + [P(A \cap B) + P(A^c \cap B)]$$

Bonferroni method

If we want to just make k comparisons, where k is small, then we can use the (trivial) "Bonferroni method": We simply use the individual t confidence interval method, adjusting the significance level (or error rate of confidence intervals) to α/k.

E.g., suppose want to make 3 comparisons. Let

$A_1 = \{$CI #1 does not contain the true value$\}$

$A_2 = \{$CI #2 does not contain the true value$\}$

$A_3 = \{$CI #3 does not contain the true value$\}$

If $P(A_1) = P(A_2) = P(A_3) = \alpha/3$, then by Bonferroni inequality

$P(A_1 \cup A_2 \cup A_3) \leq \alpha/3 + \alpha/3 + \alpha/3 = \alpha$.

So need to make 98.33% confidence intervals. for 95% familywise coverage level

Recall: A t individual CI at level $100(1-\alpha)\%$ has the following critical constant:

$$t_{1-\alpha/2, n_T - r}$$

Suppose we plan to use the Bonferroni method to estimate k parameters. Our critical constant for each of the k Bonferroni CIs:

$$t_{1-\alpha/(2k), n_T - r}$$

In R:

$r=4$, $n_T = 4 \cdot 10 = 40$, $n_T - r = 40 - 4 = 36$

Ex. Rust, all pairwise comparisons Find the Bonferroni critical constant for level $\alpha = .05$, and compare to the critical constants for the Tukey and Scheffé methods. $k =$ # comparisons $= \binom{r}{2} = \binom{4}{2} = 6$

```
> qt(1 - .05/(2*6), df=36)
[1] 2.791972
```

Tukey: $\frac{1}{\sqrt{2}} q_{.95,4,36} = 2.6932$

Bonferroni: $t_{1-.05/(2\times6),36} = 2.792$

Scheffé: $\sqrt{3F_{.95,3,36}} = 2.9324$

CAUTION: The Bonferroni method requires you to specify the parameters to be estimated in advance of the experiment. It is tempting to misapply this method in ways that lead to illegitimate "snooping."

Ex. Rust, chart of Scheffé intervals for seven comparisons. Can we use Bonferroni method? The critical constant: 2.852558

```
> qt(1 - .05/(2*7), df=36)
[1] 2.852558
```

EASY TO DO. BUT WRONG/DISHONEST HERE.

Recall 1. was contrast between 2 best & 2 worst inhibitors selected by "data snooping."

ADVANTAGE

Bonferroni applies to any parameter.