8. Single Comparisons

The F test in Section 7 is only concerned with the question of whether the treatment means are different. As we rejected the null hypothesis about the equality of the treatment means, it is natural to ask which pairs of treatment means differ significantly and by how much. The former question refers to hypothesis testing about the difference between two treatment means, the latter to confidence interval about the difference.

In order to test the hypothesis about the difference between the treatment means $\mu_i - \mu_j$, we use the two-independent sample t-test, with the important difference that the pooled estimate of standard deviation is from all groups, not just from those being compared:

$$t = \frac{\bar{x}_i - \bar{x}_j}{s_p \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}}.$$

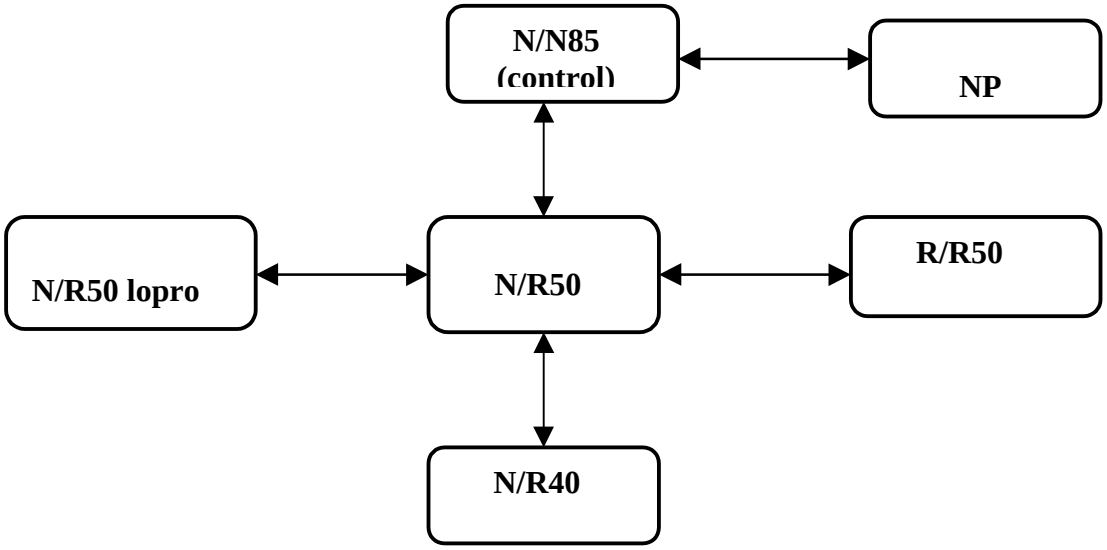
The number of groups is denoted by k, the number of observations within the ith group by n_i, and the total number of observations from all groups combined by $N=n_1+n_2+\ldots+n_k$. The values $\bar{x}_i, \bar{x}_j$ are the corresponding group means. The t statistic follows a t distribution with the number of degrees of freedom DF=N-k. The value of s_p^2 is the pooled estimate of common standard deviation from all groups and is given in the ANOVA table as the error mean square (within groups). It can also be calculated from the formula:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots(n_k - 1)s_k^2}{N - k}.$$

The fact that the pooled estimator s_p^2 from all groups is used rather than the pooled estimator from just the two groups being compared increases the power of the tests. The value of s_p^2 provided by SPSS output is 44.5989, and therefore $s_p = 6.68$.

We will use the t statistic defined above to make the comparisons among some treatment groups of interest. The structure of the planned comparisons in the study is diagrammed in the following display.

Note that comparisons other than those indicated by the arrows are not directly meaningful because the group treatments differ in more than one way. The N/N50 lopro and the N/N40 groups, for example, differ in both the protein composition and the total calories in the diet, so a difference would be difficult to attribute to a single cause.



Using this method, pairs of treatment means may be compared, and each statement will carry an observed individual level of significance α (P-value).

SPSS does not have a feature enabling you to make pairwise comparisons using the t test with the pooled estimate of standard deviation from all groups. Nevertheless, calculating the value of t from the above formula is relatively easy given the group means provided by SPSS in ANOVA output. The number of degrees of freedom for the t distribution is DF=N-k=349-6=343.

Problem	Groups compared	Hypotheses	t statistic	P-value	Conclusion
Does reducing the diet from 85 to 50 kcal/wk increase lifespan?	N/R85 N/R50	H_0: $\mu_2-\mu_4=0$ H_A: $\mu_2-\mu_4<0$	-8.08802	1.06E-14	Convincing Evidence
Does reducing the calories before weaning increase lifespan?	N/R50 R/R50	H_0: $\mu_4-\mu_5=0$ H_A: $\mu_4-\mu_5<0$	-0.49307	0.6223	No Evidence
Does lifetime on the 40 kcal/wk diet exceed the lifetime on the 50 kcal/wk diet?	N/R50 N/R40	H_0: $\mu_4-\mu_6=0$ H_A: $\mu_4-\mu_6<0$	-2.40757	0.01659	Moderate Evidence
Does reduction in protein, with same calories, decrease lifespan?	N/R50lopro N/R50	H_0: $\mu_3-\mu_4=0$ H_A: $\mu_3-\mu_4<0$	-2.18801	0.02934	Moderate Evidence
Do the control mice tend to live longer than the laboratory mice do?	NP N/N85	H_0: $\mu_1-\mu_2=0$ H_A: $\mu_1-\mu_2<0$	-4.06547	5.95E-05	Convincing Evidence

The above stated hypotheses are formulated in terms of the differences between the average responses. As our case study is an example of a randomized experiment, the hypotheses can also be worded in terms of treatment effects. The null hypothesis is that the two treatments have no differential effect. A one-sided alternative is that the average effect of a specified treatment is expected to be greater than (or less than) the average effect of the other treatment. A two-sided alternative simply states that the two treatments have different average effect; the direction of the difference is not stated.

It is possible that the experiment has detected a statistically significant difference among the treatments that is of no practical importance. Indeed, the large sample sizes (between 49 and 71) in our experiment are capable of detecting very small differences that are not of much practical importance. Thus, it is advisable to estimate the magnitudes of the differences among the average impacts of the treatments using the method of confidence intervals. Construction of a confidence interval for an additive treatment effect is precisely the same as for the difference between population means, μ_i- μ_j.

In general, we can show that a (1-α)*100% confidence interval for the difference μ_i- μ_j, $i \neq j$, is given by

$$(\bar{x}_i - \bar{x}_j) \pm t_{1-\alpha/2, N-k} \circ s_p \circ \sqrt{\frac{1}{n_i} + \frac{1}{n_j}},$$

where $t_{1-\alpha/2,\ N-k}$ is the 1-α/2 quantile from of a t distribution with N-k degrees of freedom.

SPSS does not have a feature enabling you to calculate confidence intervals based on the pooled estimate of standard deviation from all groups. Nevertheless, calculating the confidence interval from the above formula is relatively easy given the group means provided by SPSS in ANOVA output.

Problem	**Groups compared**	**Parameter of Interest**	**Point Estimate**	**95% Confidence Interval**
By how much lifetime on the 50 kcal/wk diet exceed the lifetime on the 85 kcal/wk diet?	N/R85 N/R50	$\mu_2 - \mu_4$	9.6 months	(7.3, 11.9) months
By how much reducing the calories before weaning increased lifetime?	N/R50 R/R50	$\mu_4 - \mu_5$	0.6 months	(-1.7, 2.9) months
By how much lifetime on the 40 kcal/wk diet exceed the lifetime on the 50 kcal/wk diet?	N/R50 N/R40	$\mu_4 - \mu_6$	2.8 months	(0.5, 5.1) months

By how much reduction in protein, with same calories decreased lifespan?	N/R50 lopro N/R50	$\mu_3-\mu_4$	2.6 months	(0.3, 4.9) months
By how much the diet restriction at 85kcal/week increased lifetime?	NP N/N85	$\mu_1-\mu_2$	5.3 months	(2.74, 7.85) months

The point estimate of $\mu_i-\mu_j$ is $\bar{x}_i - \bar{x}_j$, where $\bar{x}_i, \bar{x}_j$ are the corresponding group means. The confidence intervals provide interval estimates of the effects due to the treatments with 95% confidence.