Chapter 14 – Inference for Two Groups

This chapter tackles comparing population *averages* from two groups – for matched-pairs groups (section 14.1) or independent groups (sections 14.2-14.4) – using either CI's (confidence intervals) or HT's (hypothesis tests). The last example (14.6 on p.398) also gets into comparing population *percentages*, but we'll push that off until we discuss section 16.4 in a few weeks.

Testing Means for Matched Pairs – 'matched pairs' here can mean the same person or animal or object, a pair of twins, two strangers who are the same age and eat the same food, and so on – those in a 'block' are paired on the basis of something. We've actually already considered this situation in Ex. 12.5 on p. 345. We must assume that the population of *differences* (D) has a Normal distribution with unknown mean μ_d and SD σ_d, and we just use the one-sample t-test.

Example 14.1 on p. 386 – deals with retaking the SAT: here, $n_D = 5$, $\bar{d} = 20$, $s_d^2 = 500$, $t_{4,0.975} = 2.7765$, so the 95% CI for μ_d is

$$20 \pm 27.765 \text{ or } (-7.76, 47.76)$$

(This is such a wide CI since the sample SD is so large and since the sample size is small.) We're 95% confident that the true average change for all such students lies between a drop of 7.76 and an increase of 47.74. *Note that this CI contains 0.* This implies that if we want to test H_0: $\mu_d = 0$ versus H_A: $\mu_d \neq 0$ at the $\alpha = 5\%$ level, we would retain H_0. That is, there is a one-to-one correspondence between two-tailed tests and CI's provided the α levels match (as here).

→ Thus, for Example 14.2, since the null is rejected at the 5% level and it is a two-tailed test, we know that the 95% CI for μ_d does not contain 0. In this case, we conclude that there is a significant change.

Testing Equality of Two Independent Means – the annoying problem here is that the 'df' (degrees of freedom) for the test is not exactly known in some cases (depending on the relevant assumptions). It is known if and only if we can assume that the population *variances* are equal (section 14.4), but otherwise we're left to approximate with either (a) the *conservative* (14.2) or (b) *Welch's* approach (14.3).

The set-up for each of the approaches: we sample n_1 subjects from Group 1 and *independently* n_2 subjects from Group 2, and we either want to set a $(1-\alpha)*100\%$CI for $(\mu_1 - \mu_2)$ or we want to test the null hypothesis (HT) H_0: $\mu_1 = \mu_2$ at the α level. We calculate the sample statistics: $n_1, \bar{x}_1$, s_1 from Group 1 and $n_2, \bar{x}_2$, s_2 from Group 2. For the HT, we use the test statistic (TS): $t_{df} = (\bar{x}_1 - \bar{x}_2)/SE$, and for the CI, we use the interval $(\bar{x}_1 - \bar{x}_2) \pm t_{df}*SE$. Choose df and SE from this table:

Approach	df	SE
Conservative	$\min\{n_1 - 1, n_2 - 1\}$ = the minimum of $n_1 - 1$ and $n_2 - 1$	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Welch's	Use Formula (14.2) on p.394; given on Exams; used by Minitab	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Equal Variance	$n_1 + n_2 - 2$	$s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

Here, $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 - n_2 - 2}$ is the 'pooled' estimate of the common variance σ^2 ($= \sigma_1^2 = \sigma_2^2$) in the 'Equal Variance' case. Note that even though this last case gives an easy way to find the 'df', the idea of assuming equal variances may seem unreasonable in practice. Similarly, the 'Conservative' case is usually/often a little *too* conservative.

The following example is used to illustrate each of these methods – bare in mind, though, that we would use *only one* of these approaches in practice.

Example 14.3. Diets 1 (special) and 2 (standard) are tested on 10 and 20 rats respectively, and the response variable (Y) here is weight gain. Since the sample sizes are small, we need to assume that each of the respective weight gain populations for the two diets are *normally distributed* – with respective means μ_1 and μ_2 and with respective variances σ_1^2 and σ_2^2. (In fact, we need to assume normality unless *both* sample sizes are large, for example, above say 30.) Recall that a further assumption (that $\sigma_1^2 = \sigma_2^2$) is needed in for the 'Equal Variance' approach above, but not for the other two approaches. (*It is our responsibility to must understand & clearly list all assumptions at the outset!*) Finally, since neither Diet here was initially suspected of being 'better' all tests are two-sided, as are the CI's.

Conservative Approach: SE = $\sqrt{\frac{(0.80)^2}{10}+\frac{(0.60)^2}{20}}$ = 0.2864, so the TS is $t = (5.20 - 4.80)/0.2864 = 1.3969$. Since the test is two-sided, the p-value is $2 * \Pr(t_9 > 1.3969)$; note the conservative use of df = 9 here. From Table 4, $0.05 < \Pr(t_9 > 1.3939) < 0.10$, so the p-value (p) is between 0.10 and 0.20; a computer verifies that p = 0.1959. Since it exceeds α = 5%, we retain the claim that there is *no difference between the average weight gains for the two Diets*. Similarly, since the 95% CI for $(\mu_1 - \mu_2)$, $(5.20 - 4.80) \pm 2.2622 * 0.2864$ = (-0.2479 , 1.0479), *contains zero*, we reach the same conclusion using the CI.

Welch's Approach: Similar to the above, but with df = 14.24 (see p. 394 for the calculation). (The analogue of this value will be provided on exams for the relevant exam question.) Computers can deal with non-integer values for 'df' (as is the case for Minitab), but since our table requires only integer values, in order to act conservatively, we

***round this calculation down*. So, we use df = 14. The TS is still t_{14} = 1.3969. Using Table 4, still $0.05 < Pr(t_{14}>1.3939) < 0.10$, so $0.10 < p < 0.20$, but now my computer informs that P = 0.1842. Also, the 95% CI for $(\mu_1 - \mu_2)$ is now (5.20-4.80) ± 2.1448*0.2864 = (-0.2143, 1.0143); we reach the same conclusion here as above.**

Equal Variance Approach: If we have the luxury or insight of being able to in addition assume that $\sigma_1^2 = \sigma_2^2$ ($= \sigma^2$), then we estimate σ^2 by

$$s_p^2 = \frac{9*(0.80)^2 + 19*(0.60)^2}{28} = 0.45 = 0.6708^2$$

Hence, now SE = $0.6708*\sqrt{\frac{1}{10}+\frac{1}{20}}$ = 0.2598, and the TS is

$$t_{28} = (5.20-4.80)/0.2598 = 1.5396$$

Now, the p-value is $2*Pr\{t_{28} > 1.5396\}$ (again between 0.10 and 0.20 using Table 4), which is 0.1349 from a computer. Finally, the 95% CI for $(\mu_1 - \mu_2)$ is now (5.20-4.80) ± 2.0484*0.2598 = (-0.1322, 0.9322).

Even though the conclusions do not change using the above three approaches, often the results *will* change, so it is important that we use the correct methodology.