

Stat/For/Hort 571 – Midterm II, Fall 99 – Brief Solutions

(a) The observations are not paired so we carry out an independent sample T-test. $H_0: \mu_1 = \mu_2$, $H_A: \mu_1 \neq \mu_2$ and we find

$$\bar{y}_1 = 7.2, \bar{y}_2 = 6.7, s_1^2 = 0.2199, s_2^2 = 0.1933. \text{ Then}$$

$$s_p^2 = \frac{s_1^2 + s_2^2}{2} = 0.2067 \text{ and } t = \frac{(7.2 - 6.7) - 0}{\sqrt{.2067 \times (2/4)}} = 1.56 \text{ on } 6 \text{ df.}$$

The p-value = $2 \times P(T_6 > 1.56)$. We find $.10 < p < .20$. There is little evidence against H_0 and we do not reject H_0 at $\alpha = 0.10$.

(b) The assumptions are that the observations are normally distributed, independent within and between groups, and with common variance for the 2 groups.

2 (a) Compute $SS_{Error} = \sum_{i=1}^k (n_i - 1)s_i^2 = 0.04713$. Remember to

square the standard deviations given and to multiply them by sample size minus 1. Note $df_{error} = 23$ and $df_{sites} = \# \text{ sites} - 1 = 5$. Then the ANOVA table is:

source	df	SS	MS
sites	5	0.0325	0.0065
error	23	0.04713	0.00205
total	28	0.07963	

Then $F = MS_{site}/MS_{Error} = 0.0065/0.00205 = 3.17$ on 5 and 23 df. From the table we find $0.01 < p < 0.05$ and we reject at $\alpha = 0.05$, concluding that site means are not all the same.

(b) We carry out Levene's test to compare variances for 2 samples. Order each data set. The median for site 1 is 0.14 and that for site 2 is 0.18, so the absolute deviations are

site 1: 0.08 0.02 0.02 0.09
site 2: 0.03 0.01 0.00 0.02 0.03

After eliminating the single 0 deviation for site 2 (with an odd # of obsns), perform an independent sample T-test on the remaining deviations. We find

$$t = \frac{(0.0525 - 0.0225) - 0}{0.02754 \times \sqrt{2/4}} = 1.54 \text{ on } 6 \text{ df. Then } 0.10 < p < 0.20,$$

so we do not reject H_0 and we conclude that there is little evidence that variances differ.

3 (a) False. For case 1 the 95% CI is $\bar{y} \pm t_{n-1, 0.025} \frac{s}{\sqrt{n_1}}$ and for

case 2 it is $y \pm t_{4n_1-1, 0.025} \frac{s}{\sqrt{4n_1}}$. Even though $\frac{s}{\sqrt{4n_1}}$ is exactly

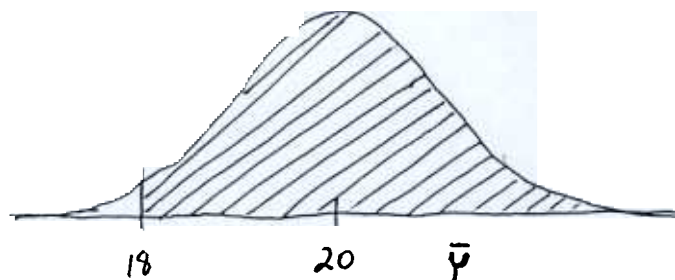
half of $\frac{s}{\sqrt{n_1}}$, the t-statistics differ because their degrees of

freedom differ, so the second interval is not exactly half as wide as the first (it will be slightly narrower).

(b) True. The observations for the 2 feet of each subject are not independent, but the subjects are and so are the means for each subject. Because the assumption of independent observations is so important for (almost) all tests, we must modify the original data so that the assumption is satisfied. Note also that the variability that matters for a comparison between the 2 groups is the variability between subjects, not the variability within subjects. Some of you noted that it might be appropriate to design a paired experiment in this situation and give each

treatment to one foot on each individual. This is a good suggestion unless there is a systemic effect of the treatment (where treating one foot affects the whole body, and thus the other foot).

4. We have $H_0: \mu = 15$, $H_A: \mu > 15$ and $Y \sim N(\mu, 70)$. We want $P(\bar{Y} > 18 | \mu = 20) = 0.95$. Graphically, we have:



$$P(\bar{Y} > 18 | \mu = 20) = P(Z > \frac{18 - 20}{\sqrt{70/n}}) = 0.95. \text{ From the table,}$$

$$P(Z > -1.645) = 0.95. \text{ So } \frac{18 - 20}{\sqrt{70/n}} = -1.645, \text{ or } \sqrt{n} = 6.88 \text{ and } n = 48.$$

5. (a) Let p be the probability of mutation. For group A, we have

$$\hat{p}_A = 66/97 = 0.6804 \text{ and for group B, } \hat{p}_B = 67/132 = 0.5076. \text{ Since the sample size is large we consider using the normal}$$

approximation. Let \hat{p} be the pooled estimate $(66+67)/(97+132) = 133/229 = 0.5808$. Then

$$Z = \frac{(.6804 - .5076) - 0}{\sqrt{.5808 \times (.4192) \times (1/97 + 1/132)}} = 2.62 \text{ and } p = 2(.0044) =$$

.009. We reject H_0 at $\alpha = 0.01$ and conclude that the proportions differ. To check the validity of the normal approximation, we check that

$$n_A \hat{p}_A = 66, n_A(1 - \hat{p}_A) = 31, n_B \hat{p}_B = 67, n_B(1 - \hat{p}_B) = 65 \text{ are all greater than } 5.$$

(b) Let $Y = \#$ horses with the mutation, $Y \sim B(n, p)$. We have $H_0: p = 0.6$ and $H_A: p > 0.6$ (one-sided). We will reject H_0 if n of n horses have the mutation and we want $P(\text{reject } H_0 | H_0 \text{ is true}) = \alpha = 0.025$. With small samples we use the binomial formula directly (note that $(1-p)n < 5$ for all n specified). Now $\alpha =$

$$P(Y=n | H_0). \text{ Then } P(Y=n) = \frac{n!}{n!} p^n (1-p)^{n-n} = p^n. \text{ We}$$

compute $p^6 = 0.0466$, $p^7 = 0.028$, $p^8 = 0.0168$, $p^9 = 0.01$, where $p = 0.6$ as specified by H_0 . So $n = 8$ is the smallest sample size for which α is less than 0.025.

Grade Distribution

100:2	
90-99:31	
80-89:39	
70-79:38	median = 79
60-69:19	
50-59:7	
<50:13	