6.8 Multiple Tests

In §6.3, we checked each other for mind-reading ability by testing $H_0 : p = .5$ vs $H_1 : p > .5$. Using level $\alpha = .05$, we found several students for whom we rejected H_0. This is _____.

If we run N tests at level α, we expect to reject H_0 when it's true for about _____ of them.

The Bonferroni Correction

The *Bonferroni correction* addresses this problem by _____. _____ is then compared to α (which is equivalent to comparing an individual P-value to _____), so the combined level is α. The Bonferroni correction is _____ in that the P-value it produces is _____ the true P-value.

Guidelines:

- If the Bonferroni-corrected P-value is _____, H_0 can be rejected _____
- If the original P-value is _____, but the Bonferroni-corrected P-value is _____, _____ the single hypothesis with small original P-value

Examples

e.g. (p. 264 #2) Five variations of a bolt-making process are run to see if any can increase the mean breaking strength of the bolts. The P-values are .13, .34, .03, .28, and .38. Which of the following choices is best?

i. Implement the process whose P-value was .03, since it performed the best.

ii. Since none had Bonferroni-adjusted P-values less than 0.05, stick with the current process.

iii. Rerun the process with $P = .03$ to see if P remains small in the absence of multiple testing.

iv. Rerun all five variations, to see if any of them produce a small P-value the second time.

e.g. (#4) Five paint additives are tested to see if any can reduce the mean drying time from the current 12 minutes. Ten specimens are painted with each type of paint, and drying times (in minutes) are measured. The results are:

	Additive				
	A	B	C	D	E
1	14.6	10.4	15.5	10.4	11.3
2	12.0	10.4	9.2	7.3	10.8
3	13.4	11.4	11.4	10.3	11.5
4	13.9	9.7	10.8	11.6	10.5
5	13.1	11.0	11.0	10.7	13.4
6	13.3	11.7	10.6	12.2	10.2
7	12.8	11.1	15.1	10.2	11.0
8	10.9	9.9	12.0	9.3	13.2
9	13.7	13.7	13.4	10.8	12.3
10	11.6	9.5	8.3	11.8	11.1
$\bar{x}$	12.9	10.9	11.7	10.5	11.5
s	1.1	1.2	2.3	1.4	1.1

For each additive, perform a hypothesis test of the null hypothesis $H_0 : \mu = 12$ against $H_1 : \mu < 12$. Assume each population is approximately normal.

a. What are the P-values for the five tests?

A. P-value $= \cdots =$ between .975 and .99

B. P-value $= \cdots =$ between .005 and .01

C. P-value $= \cdots =$ between .25 and .40

D: P-value $=$

E: P-value $= \cdots =$ between .05 and .10

b. Using significance level $\alpha = .05$, which of the following conclusions seems most appropriate?

i. At least one of the additives results in an improvement.

ii. None of the additives result in an improvement.

iii. Some of the additives may result in improvement, but the evidence is inconclusive.

___ Jelly Beans, ___ Butterfly Effect, ___ Salmon