STAtistics / Mr. Hansen Name: _____
4/15/2014 Mr. Hansen's use only: _____

Test on Chi-Square and LSRL *t* Tests

Instructions:
Each question is worth 30 points, except for the first one, which is (hopefully) a "gimme" and is worth 10 points. Full PHASTPC procedures are expected for the others, including the "S" step with the sketch labeled as "sampling distribution of _____, assuming _____" in each case. Exception: P is not required for chi-square 2-way tests. If you make a mistake, mark any work you wish to be ignored with a single large "X" to save time. Conclusions must be circled.

1. Write the digits 0 through 9 quickly and legibly. If there is evidence that you labored over the digits to make them "super-pretty," points may be deducted. The goal is to write *quickly and legibly.*

   Write the digits 0 through 9: _____

2. Three M&M's are selected at random from a production facility that produces only red and green M&M's, in the proportion of 60% red, 40% green. This procedure is repeated 250 times, and the following results are obtained:

   0 red, 3 green: 20 times
   1 red, 2 green: 70 times
   2 red, 1 green: 109 times
   3 red, 0 green: 51 times

(a) Prove that if the selection of M&M's were a true binomial procedure, then the expected counts for the four cases listed above would be 16, 72, 108, and 54, respectively. If you prove one of the middle cases (the 72 or the 108) with sufficient detail of explanation, you can say "similarly" or "ditto" and get away without showing the other three cases. If you don't know what you're doing, skip this part and move on to part (b).

(b) Is there evidence that the selection process used is *not* a true binomial procedure, based on the expected counts given in part (a)? Test at a significance level of 0.05. Use both sides of this sheet.

3. On and on it rages, the eternal debate between cake and brownies. SRS's of students from Holton, NCS, STA Lower School, and STA Upper School are asked to evaluate whether a certain dessert selection is "cake" or "brownie." The voting results are presented below.

	Holton	NCS	STA LS	STA US
"cake"	15	20	45	64
"brownie"	45	88	45	78

Are the voting proportions the same across the various populations? Test at the 0.05 level. If you find statistically significant evidence of a discrepancy, explain the largest contribution to the test statistic in context, showing your work.

4. Consider the following table showing the relationship between family size and the mean number of JetSkis owned:

family size	#JS (mean)
5	0.04399
1	0.14669
6	0.33936
3	0.40581
2	0.51470
4	0.73381
8	0.90210
9	0.99859
10	1.10855

(a) Is there a useful linear relationship between family size and the mean number of JetSkis in the household? In other words, is the slope significantly different from zero? Test at the 0.01 significance level. Show work below and on reverse.

(b) Your AP formula sheet gives a gobbledygook formula for the standard error of the slope. State the notation that is used for the standard error of the slope on the AP formula sheet. (NOT THE FORMULA, JUST THE NOTATIONAL SYMBOL.) Write the notation here: _____

(c) You probably agree that the AP formula is too messy to be useful. We will instead use a much simpler formula: standard error of slope equals $\frac{b_1}{t}$, where b_1 denotes the LSRL slope and t is the t statistic of the slope in the LSRL t test. Use this formula to compute a 95% confidence interval for the slope. Show a little bit of work, and circle your answer.

(d) Interpret the numerical value of the slope in the context of the problem. If you couldn't compute the LSRL slope, use the fake value 0.99999 when writing your answer.