

**Part I.** (50 points) Do all calculations in Minitab. Use a word processor of your choice to write a report. Insert computer text output and graphics to support what you are saying, but you need to write something that looks like an academic paper — not a pile of computer output. Turn in a hard copy of your HW in class (i.e., don't email me your HW). Also:

1. Clearly define population parameters in each problem. That is, give a verbal description of what the population mean is in the context of the problem.
2. Clearly specify hypotheses when appropriate (not every problem involves a test of hypothesis).
3. Write a coherent conclusion based on each CI or test.

- (10<sup>pts</sup>) **1. ET:** A 1997 survey conducted by M.I.T. reported that 16% of the 1,014 adults surveyed would be willing to support tax hikes to find extra-terrestrials.
- (a) (5 pts) Find a 95% CI for the proportion  $p$  of all adults that favor such a tax hike.
- (b) (5 pts) Suppose it was known that in 1990 that the proportion of all adults willing to support tax hikes to find extra-terrestrials was 0.2. Is there evidence that the proportion of adults in 1997 willing to spring for tax hikes for this purpose has changed since 1990? Carry out a test to answer this question. Use  $\alpha = 0.05$ .

10 pts

- (15<sup>pts</sup>) **2. Side effects:** Dr. Bedrick was involved in a study that examined the extent of side effects from using amphetamines to treat children with traumatic brain injuries. Prior information suggests that the probability of major side effects is very small, so it is of interest to estimate the maximum plausible value that the probability of major side effects might be. In the study none of the 15 children had major side effects.
- (a) (10 pts) Compute an exact upper 95% confidence bound for the probability of major side effects. Write a short conclusion to your analysis, interpreting the results of the exact bound in the context of the problem.
- (b) (5 pts) What happens when you have Minitab attempt to compute the bound based on the normal distribution (under [Options...]) ?

15 pts

- (25<sup>pts</sup>) **3. Suicides:** The National Center for Health Statistics (NCHS) gave the following data on the distribution of suicides in the U.S. by month in 1990. Is there any evidence that the suicide rate varies monthly, or are the data consistent with the hypothesis that the rate is constant?
- To simplify your analysis, assume the months have the same numbers of days. Compare the observed proportions across months qualitatively and through a formal goodness-of-fit test. Write a short and coherent summary to this problem. Make sure to include relevant graphical summaries with your presentation.

25 pts

Month	Num_Suicides
01Jan	1867
02Feb	1789
03Mar	1944
04Apr	2094
05May	2097
06Jun	1981
07Jul	1887
08Aug	2024
09Sep	1928
10Oct	2032
11Nov	1978
12Dec	1859

50 pts

**Part II.** (15 points)

(15<sup>pts</sup>) **1. Welsh and Breton:** The rising tide of national and regional loyalties around the world has bearing on the survival of minority or secondary languages. The article “Language Maintenance and Shift in a Breton and Welsh Sample,” (Word, 1983; p. 67–88) describes one of the first comparative studies in this area. A random sample of 86 Welsh bilingual adults yielded 76 who spoke Welsh fluently, while another random sample of 77 bilingual adults from Brittany resulted in 57 who spoke Breton fluently. Both Welsh and Breton are southern Celtic in origin. Do these data support the conclusion that the true proportion of fluent speakers among Welsh bilingual adults differs from the corresponding proportion for Breton bilingual adults?  
 To answer this question, carry out an appropriate test at the 5% level, and quantify the difference in population proportions using a 95% CI.

15 pts

(40<sup>pts</sup>) **2. Hawaiian blood:** There are four major blood groups in humans: O, A, B, and AB. A sample of individuals with records at the Blood Bank of Hawaii was selected. Each individual was classified according to blood type and ethnicity. The following two-way table of counts was obtained.

40 pts

	Ethnicity			
Blood_Type	Hawaiian	Hawaiian-White	Hawaiian-Chinese	White
O	1903	4469	2206	53759
A	2490	4671	2368	50008
B	178	606	568	16252
AB	99	236	243	5001

- (a) (15 pts) Summarize these data, focusing on comparing the proportions or percents in the 4 blood categories across the 4 ethnic groups.  
 Create a plot of each population’s proportions ( $y$ ) over blood type ( $x$ ) to help determine how the populations differ. (You can restack the data, calculate proportions, then use a time series plot with the proportions. It will take a little manual work in the spreadsheet.)
- (b) (15 pts) Is there evidence that blood type and ethnicity are associated in Hawaii? Explain.
- (c) (10 pts) Carry out any additional analyses that you deem relevant, and summarize your findings. For example, there are a number of possible additional analyses that could be done here.
- All six pairwise comparisons of blood type distributions between each pair of ethnicities.
  - Only pairs that are suggested by ethnicities: (White, Hawaiian), (White, Hawaiian-White), (Hawaiian, Hawaiian-White), and (Hawaiian, Hawaiian-Chinese).
  - There are other sensible comparisons.

Then, in each of these, if you find significant differences between blood type proportions for a pair of ethnicities, you can continue to look at individual proportion comparisons (such as a two-sample proportion test of White vs Hawaiian blood type A).

You don’t need to do tons of this. Do some amount of additional analysis to show that you understand the comparison principles. Then indicate (describe) what addition analysis *could* be done.