1. Isolation and brain-wave activity

a) I would use the analysis for two separate samples. The prisoners are not matched before random assignment; rather, the group of 20 is split into two groups without any matching.

b) \[ 0.80 - 2\sqrt{\frac{46^2}{10} + \frac{61^2}{10}}, \quad \text{to} \quad 0.80 + 2\sqrt{\frac{46^2}{10} + \frac{61^2}{10}} \]

or, after doing the math, \((.80-2(.24), .80+2(.24)) = (.32, 1.28)\)

c) We use a two-sided hypothesis test. The null hypothesis is that the two averages are equal, and the alternative hypothesis is that they differ. The test-statistic equals:

\[ t = \frac{0.80/\sqrt{\frac{46^2}{10} + \frac{61^2}{10}}}{.80/\sqrt{\frac{46^2}{10} + \frac{61^2}{10}}} = 3.31. \]

The p-value equals the area under the normal curve to the left of -3.31 and to the right of 3.31. This is a very small area, roughly .004.

Hence, we reject the null hypothesis. There does appear to be a difference in average alpha waves between confined and non-confined prisoners.

d) We assume the prisoners were randomly assigned to the treatments (they were), and that the central limit theorem holds. We would check the CLT by examining a normal probability plot to make sure the alpha wave frequencies in each group roughly follow normal curves.

e) Because the study was randomized, we can draw valid conclusions for the population of Canadian prisoners. However, I would be reluctant to extend these conclusions to populations outside of Canadian prisons. Prisoners may have different alpha-wave reactions to isolation than non-prisoners because they are different psychologically than non-prisoners.

2. Chucky Cheese

a) This is a matched pairs analysis. Pairs of plants are put in the same pot at the same time, so that they are matched. This matching invalidates the two sample analysis.

b) Using the matched pairs analysis, we get
\( (2.63 - 2\sqrt{1.97^2/15}, \ 2.63 + 2\sqrt{1.97^2/15}) \)
\[= (2.63 - 2(.51), \ 2.63 + 2(.51)) = (1.61, \ 3.65) \]

c) The null hypothesis is that the average difference equals zero. The alternative hypothesis is that the average difference (cross – self) is greater than zero.

The test statistic equals:

\[ t = \frac{2.63}{\sqrt{1.97^2/15}} = 5.15 \]

The p-value equals the area under the normal curve (actually, the proper analysis uses a t-curve with 14 degrees of freedom) to the right of 5.15, which is a very small number (less than .0001).

Hence, we reject the null hypothesis. It does appear that the height of cross-fertilized plants is larger on average than the height of self-fertilized plants.

d) We assume that the plants were assigned randomly to locations in the pot (they were), and that the central limit theorem holds. The central limit theorem holds if the differences variable follows a normal curve. We could check this using a normal probability plot in JMP.

e) It is not valid. Galton matched values based on the outcomes, not on background characteristics. By doing so, he used a wrong standard error (it’s too small). One cannot ignore the way the data are randomized.

3. A real bloody problem

a) I would use a hypothesis test for two separate samples. My null hypothesis would be that the population average blood pressures are equal in the two groups, and the alternative would be that they differ. I would need the data to follow normal curves in the two groups.

b) This is a 95% CI for a single proportion. It equals:

\[ .13 - 1.96\sqrt{\frac{.13(1-.13)}{100}} \] to \[ .13 + 1.96\sqrt{\frac{.13(1-.13)}{100}} \]
\[ = (.13 - 1.96(.036), \ .13 + 1.96(.036)) = (.06, .20) \]

The assumptions are that the data were collected at random and that the CLT holds. Because the sample size is large, the CLT holds.
c) This is a 95% CI for difference in two proportions. It equals:

\[(.13 - .10) - 1.96 \sqrt{\frac{.13(1-.13)}{100} + \frac{.10(1-.10)}{100}} \text{ to } (.13 - .10) + 1.96 \sqrt{\frac{.13(1-.13)}{100} + \frac{.10(1-.10)}{100}}\]

The assumptions are that the data were collected at random and that the CLT holds. Because the sample size is large, the CLT holds.

d) The physician is wrong. Because this was a randomized experiment, the background characteristics of the two groups should be similar. Hence, a comparison of the blood pressures after the treatments gives a measure of the effects of the treatments on blood pressures.

4. Lotteries

a) I would use the chi-squared analysis (analysis ii). These data are counts, not continuous data. The t-test assumes that the data are continuous. Plus, if you think about it, the sample average frequency for the digits has to equal 50, so that the t-test is completely meaningless.

b) Because the p-value is so large, there is a good chance of seeing these results when the null hypothesis is true. Therefore, we cannot reject the null hypothesis. The data are consistent with the hypothesis that each digit has a 1 in 10 chance of being selected.

c) We just need to make one frequency really small and the other really large. For example, we could make the frequency for eight to be 1, and the frequency for nine to be 99. This would mean that eight happens much less than 10% of the time, and nine happens much more than 10% of the time, thereby rejecting the null hypothesis. Note that the two frequencies should add to 100 since we are not changing the frequencies for other digits, and the original sum of the frequencies was 100.

5. Choose the right analysis

a) z-test or confidence interval for difference in two proportions.
b) t-test or confidence interval for difference in two means for separate samples
c) z-test or confidence interval for one proportion
d) regression