Hypothesis testing
Assignment.

1] a) $H_0 : p = 0.05$ (5% of children had autism)

$H_1 : p > 0.05$ (more than 5% children had autism)

b) The test to be conducted is a one sample test of proportions. The test is given by : z

test statistics $\hat{p}$ is observed statistics. p_0 is the null hypothesis.

c) $\hat{p} = 0.12$ $\quad p_0 = 0.05$ $\quad q_0 = P_0 \times 0.95 = 0.475$

$\sqrt{P_0 q_0 / n} = 0.01112$

$$\frac{\hat{p} - P_0}{\sqrt{P_0 q_0 / n}} = \frac{0.12 - 0.05}{0.01112} = 6.295$$

d) This lies outside the critical value of 1.6449 so we have enough evidence to reject the null hypothesis. $\hat{p} = \frac{46}{384}$ and p_0 is given as 0.05

2] a) $H_0: p = 0.2$

$H_1: p > 0.2$

b) $np_0 = 22 \times 0.2 = 4.4 < 10$

$nq_0 = 22 \times 0.8 = 17.6 > 10$

We expect more than ~~lox~~ 10 cars that failed to meet pollution control guidelines and more than 10 cars with proper emission systems

c) The sample is not large enough to continue the test.

3] a) $H_0: p = 0.44$

$H_1: p > 0.44$

b) $\hat{p} = 463/891 = 0.51964$ $\quad p_0 = 0.44$ $\quad \sqrt{p_0 q_0 / n} = 0.01663$

$$\frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{0.51964 - 0.44}{0.01663} = 4.78894$$

Since it lies outside the 5% critical value 1.6449 we can reject the null hypothesis.

c) $$\frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = 0.98 \; N(0,1)$$

$$0.98\,\hat{p} = \left[(-2.3263 \times 0.01663) + 0.44 < \hat{p} < (2.3263 \times 0.01663) + 0.44\right]$$

$CI = (0.40131, 0.47869)$

4] a) H_0: ~~$\mu_1 \neq$~~ $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

d) $\bar{x}_1 = 26.6$ $\bar{x}_2 = 13.8$ $\sigma_1 = 0.1$ $\sigma_2 = 0.5$ $n = 25$

$\sigma_1^2 = 0.01$ $\sigma_2^2 = 0.25$

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$$

$$\frac{(26.6 - 13.8) - 0}{\sqrt{\frac{0.01}{25} + \frac{0.25}{25}}} = \frac{12.8}{0.102} = 125.49$$

which tells us that the value is much higher than the critical value so we can reject the null hypothesis.

5] a) $n-1 = 51$ and we don't have $n=51$ in the t table so we will be using 50 as the degree of freedom.

b) 98% CI $= \bar{x} \pm t_{n-1} \frac{s}{\sqrt{n}} = 98.285 \pm \left(2.403 \times \frac{0.6824}{\sqrt{52}}\right)$

$= (98.06, 98.51)$

c) H_0: $\mu \leq 98.6$

H_1: $\mu \geq 98.6$

In the 98% CI we got (98.06, 98.51) which is less than 98.6 so we can say that we do not have enough evidence to reject ~~Ho~~ the null hypothesis.

6] a) $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

b) $\bar{x}_1 = 23.1$ $\quad \bar{x}_2 = 25.1$ $\quad S_p = 6.02$

-0.74288 is the statistic

c) It lies within the critical value -2.101 so we can say that the null hypothesis is true.

7] a) Attached is the excel sheet.

b) $(46.8 - 9.872) \pm 2.009 \times 7.34 \sqrt{\frac{1}{19} + \frac{1}{29}}$

$(32.576, 39.0944)$

55	30.2
45.7	2.2
43.3	7.5
50.3	4.4
45.9	22.2
53.5	16.6
43	14.5
44.2	21.4
44	3.3
33.6	10
55.1	1
48.8	4.4
50.4	1.3
37.8	8.1
60.3	6.6
46.6	7.8
47.4	10.6
44	10.6
	16.2
	14.5
	4.1
	15.8
	4.1
	2.4
	3.5
	8.5
	4.7
	18.4

	Variable 1	*Variable 2*
Mean	46.8	10.169
Variance	41.19555556	55.8858
Observations	19	29
Pooled Variance	50.13743628	
Hypothesized Mean Differ	0	
df	46	
t Stat	17.52763427	
P(T<=t) one-tail	2.72001E-22	
t Critical one-tail	1.678660414	
P(T<=t) two-tail	5.44002E-22	
t Critical two-tail	2.012895599	