Q.1 (2 points) You are reading an article that reports statistical analysis. The article contains the testing of alternative hypothesis "$\mu > 34.5$" by using a data set of size 35, and says that the test statistic is computed as $T = 2.315$.

(a) By using the critical points $t_{\alpha,34}$ for $\alpha = 0.1$, 0.05, and 0.01 ($t_{0.1,34} = 1.3070$, $t_{0.05,34} = 1.6909$, and $t_{0.01,34} = 2.4412$, respectively), find out which one of the following statements is correct for the p-value p^* of the test.

(i) $p^* \le 0.01$ (ii) $0.01 < p^* \le 0.05$ (iii) $0.05 < p^* \le 0.1$ (iv) $0.1 < p^*$

Since $t_{0.05,34} < 2.315 < t_{0.01,34}$, the correct one is (ii).

(b) Write a conclusion of the statistical analysis.

If $\alpha = 0.01$ is chosen, there is no evidence to support the claim "$\mu > 34.5$." If $\alpha = 0.05$ or 0.10 is chosen, there is evidence to support the claim "$\mu > 34.5$."

Q.2 (4 points) Company B makes olfactory sensor—a kind of electronic nose which can detect an odor with concentration of down to 200 odor units/m^2 —the threshold of current technology. Recently Company B announced that they have just invented a new type of sensor, and that they have already investigated the threshold values for their sensor with 17 different sources of odor. According to them, the sample standard deviation of threshold values is 20, suggesting that the new sensor is quite stable for various sources of odor. And the result of statistical test indicates that the mean value of threshold is less than 200 odor units/m^2. They even said "the p-value was 0.022 for the test." However, they are disclosing neither the actual data set nor the sample mean.

(a) What are their null and alternative hypothesis for the mean threshold value μ for the new sensor?

$H_0 : \mu = 200$ vs. $H_A : \mu < 200$.

(b) With an unknown variable $\bar{X}$ of the sample mean, write down the test-statistic T used by Company B.

$$T = \frac{\bar{X} - 200}{20/\sqrt{17}}$$

(c) The following table shows the critical point $t_{\alpha,df}$ of t-distribution for various values of n and α.

	α	
df	0.011	0.022
15	2.555	2.199
16	2.536	2.186
17	2.521	2.175
18	2.507	2.166
19	2.494	2.157

Then, can you find the value of T which Company B actually calculated for their statistical test?

$T = -t_{0.022,16} = -2.186$

(d) Find the sample mean $\bar{X}$ which Company B refused to disclose.

$\bar{X} = 200 - (2.186)\left(\frac{20}{\sqrt{17}}\right) \approx 189.4$

Q.3 (4 points) In a test for a particular disease, it was assumed that *false-positive* results are obtained about 1 in 100. When the test is administered to 10,000 healthy people (who are known to be free of the disease), 122 false-positive results were observed.

(a) Assuming that false-positive results are obtained about 1 in 100, what is the *approximate* distribution for the number X of false-positive results? What are the parameters?

X has approximately a normal distribution with mean $np = 100$ and variance $np(1-p) = 99$.

(b) Assuming that false-positive results are obtained about 1 in 100. Estimate *approximately* the probability that we obtain 122 false-positive results or more. You may use $\sqrt{99} \approx 10$.

$P(X \geq 122) \approx 1 - \Phi(2.2) \approx 0.014$.

(c) Can you find statistical evidence that there are more than one false-positive result for every 100 tests? Starting from the statement of hypotheses, justify your answer by completing the hypothesis test.

We test $H_0 : p = 0.01$ vs. $H_A : p > 0.01$. Note that what we have calculated in (b) is the p-value. Thus, if you choose the significance level $\alpha = 0.05$ or 0.1, then we reject H_0 and we find evidence that there are more than one false-positive result for every 100 tests.

If $\alpha = 0.01$ then we fail to reject H_0 and there is nothing to say about evidence.