Student Information

Name :Alperen Oğuz Çakmak
ID :2237162

Answer 1

a)

Firstly, we know that these samples are collected independently of each other. Stating that X is samples from people with age 40 and above, and Y is samples from people with age below 40. We want to compute confidence interval of the difference between the means such as $\theta = \mu_x - \mu_y$.

Proposed estimator is $\hat{\theta} = \bar{X} - \bar{Y}$ since $\bar{X}$ estimates μ_x and $\bar{Y}$ estimates μ_y.

Question gives us both sample means and sample standard deviations for X and Y. But the standard deviation is unknown and the sample sizes for X and Y is small sizes. Because of that we need to use student's t distribution to obtain the confidence interval of differences of means. And since population variances are to be assumed equal, we will use formula 9.11 (from textbook page 261) to obtain pooled standard deviation. Sample sizes n and m are 19 and 15 respectively.

$s_p^2 = \frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2} = \frac{18\times 0.921+14\times 1.124}{19+15-2} = 1.06$ and $s_p = 1.033$

Formula for confidence interval for the difference between means with equal but unknown standard deviation is $\bar{X} - \bar{Y} \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$. Since this part of question asks for 95% confidence interval, (1 - α) = 0.95 and $\alpha = 0.05$ so we need $t_{0.025}$ with degrees of freedom 32 (n+m-2) which is equal to 2.037 (from table A5)

Now we have all of the values of formula and we can compute confidence interval.

$\bar{X} - \bar{Y} = 3.375 - 2.05 = 1.325$
$t_{\alpha/2} s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = (2.037) \times (1.033) \times \sqrt{\frac{1}{19} + \frac{1}{15}} = 0.726$

So our confidence interval is 1.325 ± 0.726 or [0.598, 2.051].

b)

This part is same with previous one except it asks for a 90% confidence interval. So we know that $(1-\alpha) = 0.9$ and $\alpha = 0.1$ so we need to know $t_{0.05}$ with degrees of freedom 32. It is equal to 1.694 (from table A5). All the other values are the same so we can directly replace $t_{0.05}$ value and compute.

$\bar{X} - \bar{Y} = 3.375 - 2.05 = 1.325$
$t_{\alpha/2} s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = (1.694) \times 1.033 \times \sqrt{\frac{1}{19} + \frac{1}{15}} = 0.6$

So our confidence interval is 1.325 ± 0.6 or [0.725, 1.925].

c)

Since the sample size is small and standard deviation is unknown for people with age 40 and above we will use confidence interval for unknown σ formula (9.9) that is $\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ and $\alpha = 0.05$ because of 95% confidence interval. It can be seen that $t_{0.025} = 2.101$ (from table A5) with degrees of freedom is n - 1 = 19 - 1 = 18. Calculating the result, corresponding margin is 0.462 and 95% confidence interval for people with age 40 and above is interval [2.912, 3.837] but this interval slightly exceeds the lower bound of supporting BREXIT because 3 means neutral and our interval includes the interval [2.912, 3.0] which we cannot say people with age 40 and above support BREXIT.

In order to find an interval which excludes [0,3] we would need to change margin that has a formula $t_{\alpha/2} \frac{s}{\sqrt{n}}$ and in order to decrease margin, t value could be decreased (since standard error will not change because we don't want to change our sample). Decrease in the t value will increase α and because of confidence level being $(1-\alpha)100\%$ our confidence level would be lower than 95%.

Hence, the answer is no. We cannot say people with age 40 and above supports BREXIT with 95% confidence level.

Answer 2

a)

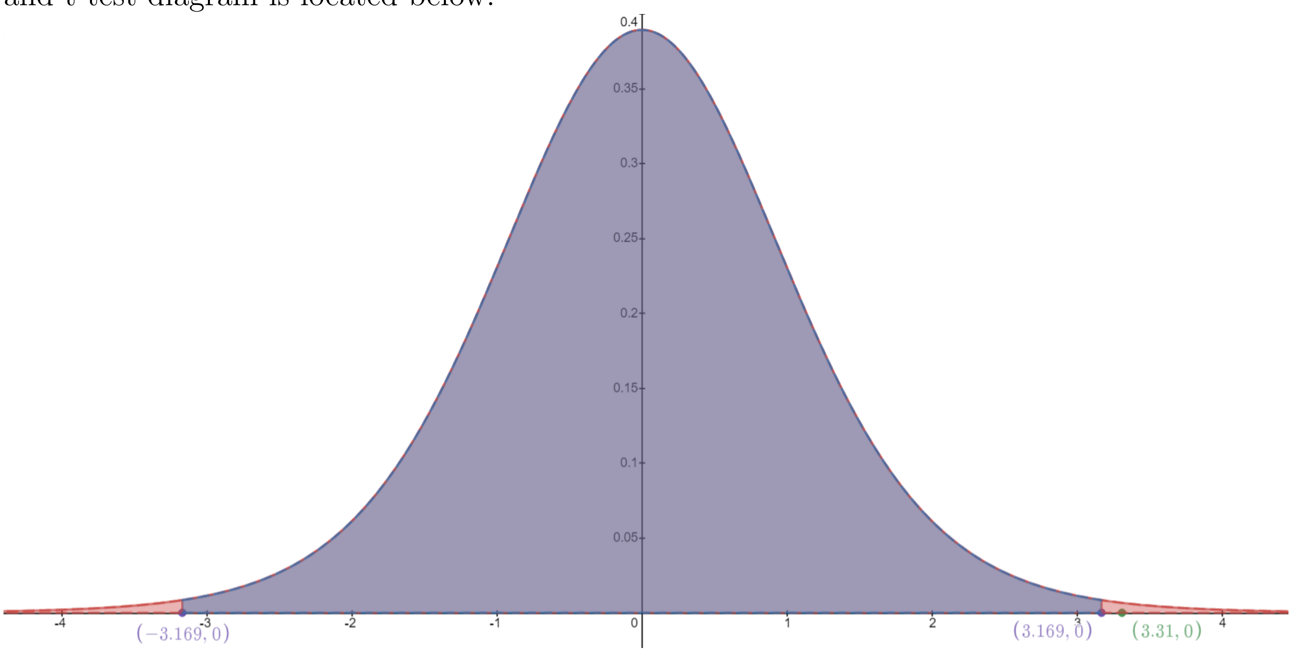
Since a product is qualified if it weights 20 kilograms so:
$H_0 : \mu = 20$

b)

Any product which is not 20 kilograms will be disqualified so:
s $H_A : \mu \neq 20$

c)

Since standard deviation is unknown we will apply t-test with formula $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$ in order to find test statistic t. The question tells us $\bar{X} = 20.7$, $\mu_0 = 20$, $s = 0.07$, $n = 11$ and $\alpha = 0.01$. Calculating the test statistic we find that t = 3.316 now we have to find the rejection region. Since the alternative hypothesis is two-sided, we need $t_{\alpha/2} = t_{0.005}$ which can be derived from t-table with n-1 = 10 degrees of freedom. And that corresponds to 3.169 . Because of two-sided alternative case (9.20) $|t| \geq t_{\alpha/2}$, null hypothesis is rejected. The rejection region is $(-\infty, -3.169] \cup [3.169, \infty)$ and t-test diagram is located below.



Two red dots located between borders for red and blue regions are points (-3.169,0) and (3.169,0) which are critical values. The areas under the curve divided into 2 regions. Red is the rejection region which includes interval $(-\infty, -3.169] \cup [3.169, \infty)$ under the curve. Our test statistic t is 3.316 which resides in rejection region. Thus, null hypothesis is rejected. In other words, they should stop producing and check the line.

Answer 3

a)

Considered new pain killers as X and the current painkiller as Y, The new painkiller should have less mean than the current one in order to be considered better. So the difference between their mean values should be more than 0, but the null hypothesis can't be an inequality so we can say that the difference is 0 and they are equal, so the null hypothesis is:
$H_0 : \mu_x = \mu_y$

b)

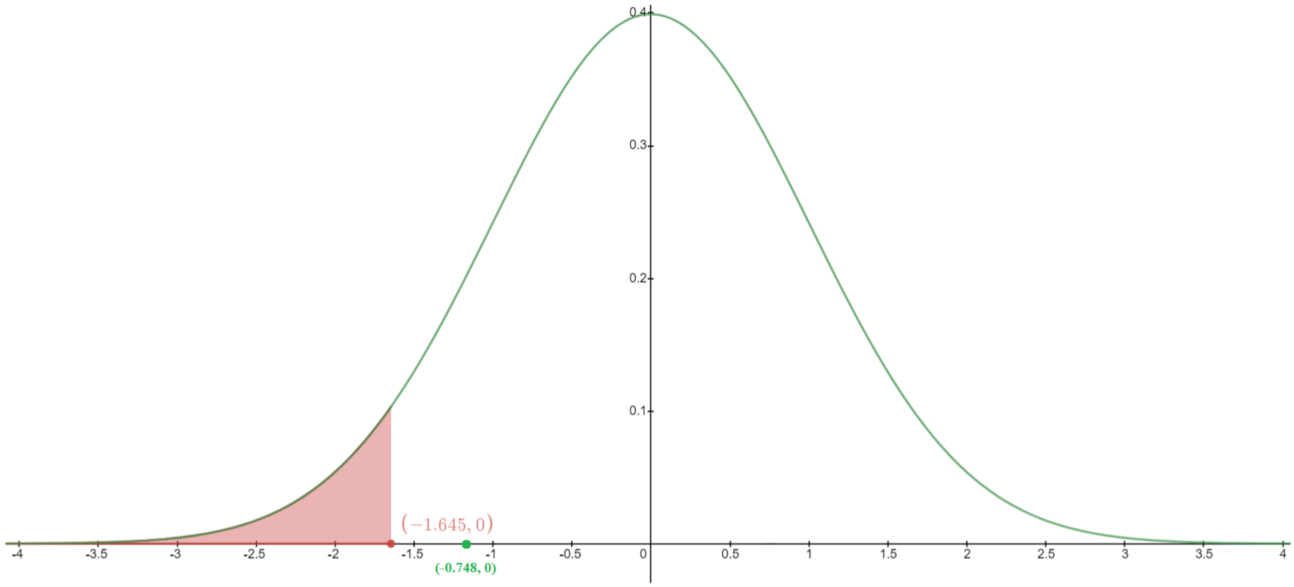
If the new painkiller takes more time than the current one it will be considered that the new one is not better than the current one so new ones should have less mean than the current ones if they were to be considered better. So my choose of alternate hypothesis is:
$H_A : \mu_x < \mu_y$

c)

Standard deviations of these two samples are unknown but the sample size is bigger than 30 so we can use normally distributed z-test in this question. Since we want 5% level of signification and the test is one-sided, alternate hypothesis is an inequality, $z_\alpha = z_{0.05}$ that is 1.645 which is a common z value.

We are given that $\bar{Y} = 3$, $\bar{X} = 2.8$, $s_y = 1,4$, $s_x = 1.7$ and n=68. We can compute test statistic z with the two-sample Z-test formula (Table 9.1: Summary of Z-tests.) replacing sigmas with our sample standard deviations since sample size is large enough.
Doing the computations, we find z = $\frac{\bar{X}-\bar{Y}-D}{\sqrt{\frac{s_x^2}{n}+\frac{s_y^2}{n}}}$ = -0.748.

Since $z_{0.05} = 1.645$ and because of alternate hypothesis, our test is left-tail alternative. So our rejection region is $(-\infty, -1.645]$ and our test statistic z is computed as -0.748 which is not in the rejection region. Thus, we accept the null hypothesis and also rejected the alternate hypothesis which actually states that the new pain killers takes less time than the old ones. Hence, it can be clearly seen that the new painkillers are not better than the current ones in the market.



Red region is the rejection region and the green dot is where our test statistic resides.