Answer on Question #42549 – Math – Statistics and Probability

A town-planning sub-committee in Tshwane wanted to know if there is any difference in the mean travelling time to work of car and Train commuters. They there carried out a survey amongst car and bus commuters and with the following sample statistics:
Car Commuters
Train Commuters
X1= 29.6 min
X2 = 25.2 min
S1= 5.2 min
S2= 2.8 min
N1=22 drivers
N2=36 passengers
4.1 Test the hypothesis at the 5% significance level that it takes car commuters to get to work earlier than Train commuters .

Solution

Case 1. Variances are not assumed to be equal.

Assumptions: normal, independent samples; σ_1^2, σ_2^2 are unknown, $\sigma_1^2 \neq \sigma_2^2$; n_1 and n_2 are small

The null hypothesis $H_0: \mu_1 = \mu_2$ versus $H_a: \mu_1 < \mu_2$. Looking at the observed sample standard deviations, we note that s_1 is approximately twice s_2 so the assumption $\sigma_1 = \sigma_2$ is suspect.

We therefore use the conservative test based on the test statistic:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{X_1 - X_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} = \frac{4.4}{1{,}203} \approx 3.66.$$

In some books degrees of freedom d.f.=smaller of $n_1 - 1$ and $n_2 - 1 = 21$.
For d.f.=21, the tabled value is $t_\alpha = t_{.05} = 1.72$. We set the rejection region $R: T < -1.72$. The observed value of the test statistic is $T^* = 3.66$ and it is not in R. Therefore, at the $\alpha = .05$ level of significance the null hypothesis H_0 is accepted and we conclude that the claim (it takes car commuters to get to work earlier than Train commuters) is not substantiated by the data.

In some books degrees of freedom is calculated d.f.= $\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \approx 28.$

Case 2. Variances are assumed to be equal.

Assumptions: normal, independent samples; σ_1^2, σ_2^2 are unknown, $\sigma_1^2 = \sigma_2^2$; n_1 and n_2 are small.

As a working rule, the range of values $\frac{1}{2} \leq \frac{s_1}{s_2} \leq 2$ may be taken as reasonable cases for making the assumption $\sigma_1 = \sigma_2$. In this problem $\frac{s_1}{s_2} = 1.85$.

The null hypothesis $H_0: \mu_1 = \mu_2$ versus $H_a: \mu_1 < \mu_2$.

We use the test based on the test statistic:

$$T = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-1}} \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{X_1-X_2}{\sqrt{\frac{(N_1-1)s_1^2+(N_2-1)s_2^2}{N_1+N_2-2}} \times \sqrt{\frac{1}{N_1}+\frac{1}{N_2}}} \approx 4.23.$$

In some books degrees of freedom d.f.=smaller of $n_1 - 1$ and $n_2 - 1 = 21$.
For d.f.=21, the tabled value is $t_\alpha = t_{.05} = 1.72$. We set the rejection region $R: T < -1.72$. The observed value of the test statistic is $T^* = 3.66$ and it is not in R. Therefore, at the $\alpha = .05$ level of significance the null hypothesis H_0 is accepted and we conclude that the claim (it takes car commuters to get to work earlier than Train commuters) is not substantiated by the data.

In some books degrees of freedom is calculated d.f.=$n_1 + n_2 - 2 \approx 56$.