DANA 4800 Homework #5

A local pizza joint has recently hired additional drivers and claims that its average delivery time for orders is under 25 minutes. A random sample of 20 customer deliveries was examined and the average delivery time was found to be 22.3 minutes with a standard deviation of 5.6 minutes. The delivery time generally follows a Normal distribution.

Note: Please use t-distribution to complete this question.

a) Identify the parameter of interest. [2 marks]

Subjects = customers

Variable = deliver time

Type = numerical time

Define μ as the average of deliver time in minutes, of all customer deliveries from local pizza

b) Set up the null hypothesis and the alternative hypothesis. [2 marks]

H0: mu ≥ 25 min

Ha: mu < 25 min

c) Calculate the test statistic and find the p-value. [1+3]

$$TS = \frac{\bar{X} - \mu 0}{s/\sqrt{n}} = \frac{22.3 - 25}{5.6/\sqrt{20}} = -2.156$$

$$v = n - 1 = 20 - 1 = 19$$

99.9%

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Free Usage Disclaimer: Feel free to use and share the above images of T-Table as long as you provide

In row 19 of the t-table, the absolute value of the TS value falls between 2.0930 (with 0.025 tail prob) and 2.5395 (with 0.01 tail prob).
P-value is between 0.01 and 0.025.

Could be a **middle point** as a suggestion but in somehow is better to say is between a range 0.01 -0.025

$$\text{Midpoint} = \frac{0.01 + 0.025}{2} = \frac{0.035}{2} = 0.0175$$

d) Draw an appropriate conclusion using 5% significance level. [2 marks]

T.S. <5%

Since the p-value is less than the 5% level of significance, we have enough statistical evidence to **reject the null** hypothesis and conclude that the average delivery time (in minutes) for all deliveries to customers at the local pizza restaurant is **significantly** smaller from 25 min.

2. What is the effect of concussions on the brain? Researchers measured the brain sizes (Hippocampal volume in litre) of 25 college football players with a history of clinically diagnosed concussion and 25 college football players without a history of concussion. Is there evidence of a difference in mean brain size between football players with a history of concussion and those without concussions? It is believed that the brain size follows a Normal distribution and the equal variance assumption is valid in this question. The following table shows the summary statistics.

Group	Sample Size	Sample Mean	Sample S.D.
Concussion	25	5.7	0.65
Non-concussion	25	6.3	0.85

a) Identify the parameter of interest. Make sure you clearly identify which one is Population 1 and which is Population 2. [2 marks]
Subjects = college football players
Variable = μ brain size each group, compare
Type = numerical value

Groups:
Population 1 = players with a history of concussion ($\boldsymbol{\mu 1}$)
Population 2 = players without a history of concussion ($\boldsymbol{\mu}2$)

Define $\boldsymbol{\mu 1} - \boldsymbol{\mu 2}$ as the difference in average from brain size each group

b) Set up the null hypothesis and the alternative hypothesis. [2 marks]
$\boldsymbol{H0}$: $\boldsymbol{\mu 1 - \mu 2 = 0}$ vs.
$\boldsymbol{Ha}$: $\boldsymbol{\mu 1 - \mu 2 \neq 0}$

c) Calculate the test statistic (with equal variance assumption). [2 marks]
Equal variance assumption or (if the ratio of the two variances is between 0.5 and 2)

s1/2=0.65^2=0.4225
s2/2= 0.85^2=0.7225

$$\frac{0.7225}{0.4225} \approx 1.71$$

This ratio is **less than 2**, so we **can assume equal variance is valid**.

$$\mathrm{S2/p} = \frac{\frac{(\mathrm{n1}-1)\mathrm{s1}}{2} + \frac{(\mathrm{n2}-1)\mathrm{s2}}{2}}{\mathrm{n1}+\mathrm{n2}-2} = \frac{24(0.4225) + 24(0.7225)}{25+25-2}$$

$$\frac{\mathrm{S2}}{\mathrm{p}} = \frac{10.14 + 17.34}{48} = 27.48 = 0.5725$$

$$sp = \sqrt{0.5725} \approx 0.7566$$

$$\mathrm{SE} = \mathrm{sp} \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}} = 0.7566 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}} = 0.7566 \cdot \sqrt{0.08} = 0.7566 \cdot 0.2828 \approx 0.2141$$

V=DF=25+25-2=48

$$\mathrm{t} = \frac{\overline{\mathrm{X1}} - \overline{\mathrm{X2}}}{SE} = \frac{5.7 - 6.3}{0.2141} = \frac{-0.6}{0.2141} \approx -2.80$$

$$\text{This is a two} - \ \text{Ha:}\, \mu 1 - \mu 2 \neq 0$$

d) Determine the size of the p-value, using the t-table. [2 marks]
Step 2: Use the t-table

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
					Confidence Level						

48

From a t-table:
For DF = 40–60:
At-value of 2.704 gives a two-tailed p ≈ 0.01
At-value of 3.307 gives a two-tailed p ≈ 0.002

e) Use 1% level of significance and draw a conclusion. [2 marks]

Since the p-value is **smaller than the 1% significance level**, we have enough statistical evidence to **reject the null hypothesis** and conclude that the difference in average brain size between college football players **with and without a history of concussion** is **significantly different from zero**. In other words, the average hippocampal volume between the two groups of players is **significantly different**.

3. A toothpaste manufacturer claims that children brushing their teeth daily with this company's new toothpaste product will have fewer cavities than children using a competitor's brand. In a carefully supervised study where children were randomly assigned to one of the two brands of toothpaste for a two-year period, the number of cavities for children using the new brand was compared with the number of cavities for children using the competitor brand. You can assume that the number of cavities follows a Normal distribution and the equal variance assumption is not valid. The data set "DANA4800_HW5_Data.xlsx" contains the results and can be downloaded from Brightspace.

.

a) Define the parameter of interest. Make sure you clearly identify which one is Population 1 and Which is Population 2. [2 marks]

Define the parameter of interest. [2 marks]
Subjects: Children using toothpaste

Variable of Interest: Number of cavities (numerical)

Population 1: Children using the new toothpaste (μ_1)
Population 2: Children using the competitor's brand (μ_2)

Parameter of interest: The difference in mean number of cavities between the two groups:

b) Set up the null hypothesis and the alternative hypothesis. [2 marks]

The manufacturer claims that $\mathbf{\mu_1 < \mu_2}$ (fewer cavities with the new toothpaste).

- **Null hypothesis (H_0)**: $\mu1-\mu2\geq0$
- **Alternative hypothesis (H_a)**: $\mu1-\mu2<0$

c) Use t.test() in R to find the test statistic and p-value. [2 marks]

```
library(readxl)

path<-"P:/langara/term 1/DANA-4800-001 - Data Analysis and Stat
Infer  20287.202520"
file <- file.path(path, "DANA4800_HW5_Q3_Data.xlsx")
group <- read_excel(file)

# Create a frequency table of the variable 'Distracted'
freq_table <-table(group$Distracted)


t.test(group$New,group$Competitor,
       alternative="less",
       var.equal = FALSE)
```

```
-----------------------------------------------------------------
data:  group$New and group$Competitor
t = -0.83075, df = 13.547, p-value = 0.2103
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf 0.7160158
sample estimates:
mean of x mean of y
 2.000000  2.636364
```

d) Identify the sampling distribution of the test statistic. [2 marks]

The test statistic follows a t-distribution with approximately 17.7 degrees of freedom, which does not assume equal variances.

Variance (New toothpaste) = 0.727
Variance (Competitor toothpaste) = 4.145

d) Test the manufacturer's claim using 10% significance level. [2 marks]

Since the p-value is bigger than the 10% significance level, we don't have enough statistical evidence to reject the **null hypothesis** and conclude Therefore, we do not have sufficient statistical evidence to support the manufacturer's claim that the new toothpaste results in fewer cavities compared to the competitor's brand.