P382 10-8 (Modified) (9 pt)

A polymer is manufactured in a batch chemical process. Viscosity measurements are made on each batch, and long experience with the process has indicated that the variability in the process is fairly stable with σ = 20. Fifteen random batch of viscosity measurements yields a sample mean of 750.2. A process change that involves switching the type of catalyst used in the process is made. This change does not change the variability in the process. Following the process change, eight (independent) batch viscosity measurements are taken at random that has generated a sample mean of 756.88.

a. If $\mu_1 - \mu_2$ is the mean batch viscosity before and after the process change, can you conclude that $\mu_1 - \mu_2$ is less than 10 at 0.10 level of significance? Use the confidence interval method – that is, the solution will include a confidence interval, a final conclusion and an explanation for how this interval is used to draw your conclusion.

$H_0 : \bar{x}_1 - \bar{x}_2 = 10$

$H_1 : \bar{x}_1 - \bar{x}_2 < 10$

$(-\infty, \bar{x}_1 - \bar{x}_2 + z_{0.1}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}]$

$\rightarrow (-\infty, 750.2 - 756.88 + 1.28\sqrt{\frac{20^2}{15} + \frac{20^2}{8}}]$

$\rightarrow (-\infty, -6.68 + 1.28\sqrt{76.667}]$

$\rightarrow (-\infty, -6.68 + 11.2076]$

$\rightarrow (-\infty, 4.5276]$

At the 90% confidence level (0.1 significance level) $\bar{x}_1 - \bar{x}_2$ is in the range $(-\infty, 4.5276]$ so therefore $\bar{x}_1 - \bar{x}_2$ is less than 10 at the 0.1 level of significance.

b. Is your method used in part a justified? Why or why not? If not, do you need any additional assumptions?

We need the following to be true:

1. Population variances are both known
2. Both populations are normal or $n_1, n_2 \geq 30$
3. Two random samples are independent

4. $n_1/N_1, n_2/N_2 \leq 5$ % in case of sampling without replacement

We know that $\sigma_1 = \sigma_2 = 20$ and samples are independent but the rest of them need to be assumed.

Problem 2 (26 pt)

The supervisor of a factory believes that method 1 takes workers lesser time to perform a task than method 2, on the average. Two independent random samples of 10 workers of the factory are selected and workers from different samples are asked to perform the task using different methods. The times used by the 20 workers are given below in seconds:

Worker	1	2	3	4	5	6	7	8	9	10
Method 1	47	44	43	49	48	44.5	48	47	46	50
Method 2	45	43	42	45	46	45	44	45	44	47

a. If we assume that both population distributions are normal with unknown but equal variances, then at a level of significance of 0.01, can you conclude that the supervisor's belief is true? Use the 8-step rejection region approach.

$H_0 : \bar{x}_1 - \bar{x}_2 = 0$

$H_1 : \bar{x}_1 - \bar{x}_2 < 0$

$df = n_1 + n_2 - 2 = 18$

$t < -t_\alpha = -2.552$

$S_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2)s_2^2}{df}}$

Test Statistic: $t = \frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{S_p\sqrt{1/n_1+1/n_2}}$

```
m1 = (c(47 ,44 ,43 ,49 ,48 ,44.5 ,48 ,47 ,46 ,50))
s1=sd(m1)
x1=mean(m1)
print(s1)
```

```
[1] 2.261391
```

```
print(x1)
```

```
[1] 46.65
```

```
m2= (c(45 ,43 ,42 ,45 ,46 ,45 ,44 ,45 ,44 ,47))
s2=sd(m2)
x2=mean(m2)
print(s2)
```

```
[1] 1.429841
```

```
print(x2)
```

```
[1] 44.6
```

$\bar{x}_1 = 46.65, s_1 = 2.261391$

$\bar{x}_2 = 44.6, s_2 = 1.429841$

$S_p = \sqrt{\frac{(9)(2.261391)^2+(9)(1.429841)^2}{18}} = 1.8918687$

$t = \frac{46.65-44.6-(0)}{1.8918687\sqrt{1/10+1/10}} = \frac{2.05}{0.8460694} = 2.422969$

$t = 2.422969$ is not less than -2.552 so we do not reject the null hypothesis.

Therefore there is not enough evidence to say ath the 0.01 level of significance that the supervisor's belief is true and that there is not enough evidence to conclude that method 1 takes less time on average than method 2.

b. Use R to conduct the same test as in part a. Attach the code and output.

```
t.test(m1,m2, mu=0, alternative="less", var.equal=TRUE, conf.level=0.99)
```

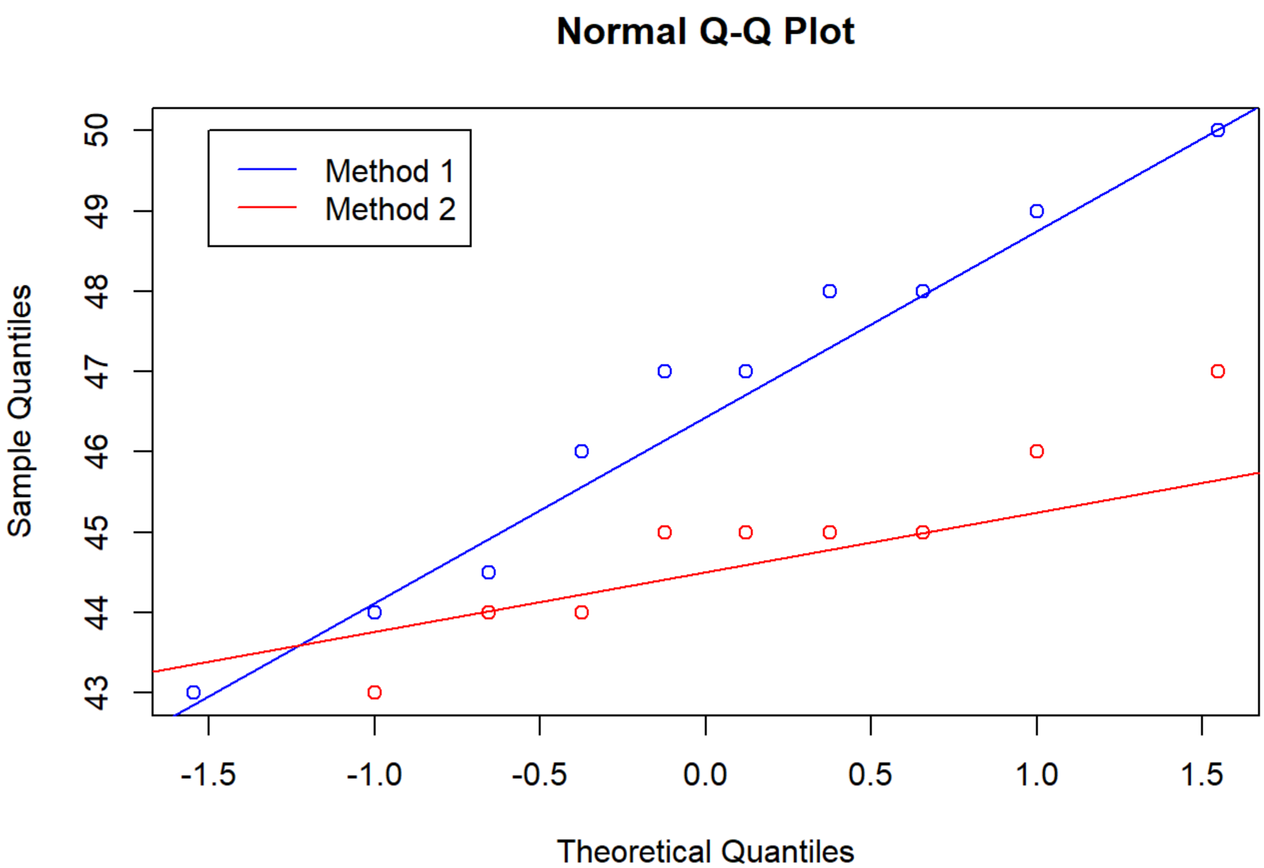
```
    Two Sample t-test

data:  m1 and m2
t = 2.423, df = 18, p-value = 0.9869
alternative hypothesis: true difference in means is less than 0
99 percent confidence interval:
    -Inf 4.20949
sample estimates:
mean of x mean of y
    46.65     44.60
```

c. Use R and the sample data to draw 2 normal probability plots in the same picture. Attach the code and output.

```
q1 = qqnorm(m1, col="blue")
q2 = qqnorm(m2, plot.it= FALSE)

points(q2, col="red")
qqline(m1, col="blue")
qqline(m2, col="red")
legend(-1.5,50, legend=c("Method 1", "Method 2"), col=c("Blue", "Red"), lwd=1, lty=c(1,1))
```



d. From the plots in part c, can you justify the assumption in part a that both population distributions are normal? Why or why not?

It is clear that method 1 is normally distributed as the points cluster around the line. However while the argument could be made that method 2 is not normally distributed, I still believe method 2 is normally distributed and has an outlier.

e. From the picture in part c, can you justify the assumption in part a that population variances are equal? Why or why not?

No because the two lines have different slopes indicating the standard deviations of the two populations are different.