CIMT Further Statistics p84 Ex 4.7

2.

X = units used by exchange X

$X \sim N(\mu_x, 100^2)$

$n_x = 15$

$\bar{x} = \frac{1}{15} \times 7980 = 532$

Y = units used by exchange Y

$Y \sim N(\mu_Y, 100^2)$

$n_Y = 20$

$\bar{y} = \frac{1}{20} \times 10220 = 511$

$H_0: \mu_X = \mu_Y$

$H_1: \mu_X \neq \mu_Y$

Assume H_0 to be true

$\alpha = 5\%$

two tailed test

we perform a 2-sample z-test as the standard deviations of the populations are known.

so $X \sim N(\mu_X, 100^2)$ $\quad$ $Y \sim N(\mu_Y, 100^2)$

$\bar{X} \sim N(\mu_X, \frac{100^2}{15})$ $\quad$ $\bar{Y} \sim N(\mu_Y, \frac{100^2}{20})$

$$\bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, \frac{100^2}{15} + \frac{100^2}{20})$$

so $$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{100^2}{15} + \frac{100^2}{20}}} \sim N(0, 1^2)$$

so $$\text{p-value} = 2 \times P\left(Z > \frac{532 - 511 - (0)}{\sqrt{\frac{100^2}{15} + \frac{100^2}{20}}}\right)$$

$= 2 \times P(Z > 0.614817)$

$= 2 \times 0.269338$

$= 0.538675$

> 0.05

so we have no evidence to reject H_0

we do not have evidence to suggest that mean units used at each exchange is not equal

3. a) a paired test is inappropriate as the 6 stores in Northtown are not associated in any way with the 6 stores in Southville. They are independent samples, not linked in any way.

b) under the stated assumptions

X_S = costs in southville $\qquad$ X_N = cost in Nortown

$X_S \sim N(\mu_S, \sigma^2)$ $\qquad$ $X_N \sim N(\mu_N, \sigma^2)$

$n_S = 6$ $\qquad$ $n_N = 6$

$H_0: \mu_S = \mu_N$

$H_1: \mu_S > \mu_N$

Assume H_0 to be true

$\alpha = 5\%$

one tail test

as variance of parent populations are unknown, we estimate them and use a 2-sample pooled t-test

we have $\bar{x}_S = 12.6$ $\qquad$ $\bar{x}_N = 12.04$

$s_S = 0.359054$ $\qquad$ $s_N = 0.384083$

$$\text{pooled, } s^2 = \frac{(6-1)\times 0.359054^2 + (6-1)\times 0.384083^2}{5+5}$$

$$= 0.37178^2$$

$$\text{test statistic, } t = \frac{12.6 - 12.04 - (0)}{s\sqrt{\frac{1}{6}+\frac{1}{6}}} \sim t_{10}$$

$$= 2.60893$$

$$\text{p-value} = P(t_{10} > 2.60893)$$

$$= 0.013044$$

$$< 0.05$$

so we have evidence to reject H_0 and conclude that the mean costs are higher in Southville than Nortown.

4.

let X_A = percentage gain for group A X_B = percentage gain for group B

$X_A \sim N(\mu_A, 4.5^2)$ $X_B \sim N(\mu_B, 4.5^2)$

$n_A = 7$ $n_B = 10.$

$H_0 : \mu_A = \mu_B$

$H_1 : \mu_A > \mu_B$ (high nitrate retards gain)

Assume H_0 to be true

$\alpha = 1\%$

one tail test

as standard deviations are known, we perform a two sample z-test for difference in population means.

$X_A \sim N(\mu_A, 4.5^2)$ $X_B \sim N(\mu_B, 4.5^2)$

$\bar{X}_A \sim N(\mu_A, \frac{4.5^2}{7})$ $\bar{X}_B \sim N(\mu_B, \frac{4.5^2}{10})$

$$\bar{X}_A - \bar{X}_B \sim N\left(\mu_A - \mu_B, \frac{4.5^2}{7} + \frac{4.5^2}{10}\right)$$

$$\frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{4.5^2}{7} + \frac{4.5^2}{10}}} \sim N(0, 1^2)$$

we have $\bar{x}_A = 19.9$ $\bar{x}_B = 15.6$

$$\text{p-value} = P\left(Z > \frac{19.9 - 15.6 - (0)}{\sqrt{\frac{4.5^2}{7} + \frac{4.5^2}{10}}}\right)$$

$= P(Z > 1.93901)$

$= 0.02625$

> 0.01

so we do not have evidence to reject H_0. We do not have evidence to suggest that the mean percentage gain for group A is greater than that of Group B.

If the mice in group B were appreciably heavier at the outset then the percentage gains may be affected by this factor. This, in turn, would alter the mean gain ($\bar{x}_B$) which would alter the value of the test statistic and possibly the conclusion of the hypothesis test at the 1% level.

5.

subject	1	2	3	4	5	6	7	8	9	10
N, no alcohol	260	565	900	630	280	365	400	735	430	900
A, alcohol	185	375	310	240	215	420	405	205	255	900
N − A	75	190	590	390	65	−55	−5	530	175	0

let X = difference in time until useful consciousness ended (where difference = no alcohol − alcohol)

we shall assume X are normally distributed

$X \sim N(\mu, \sigma^2)$

$H_0: \mu = 0$

$H_1: \mu > 0$ (ie. alcohol reduces time of useful consciousness)

Assume H_0 to be true

$\alpha = 5\%$

one tail test

as we will estimate σ, we shall use a t-test on the single sample of 10 differences.

we estimate σ with $s_{n-1} = 230.584$, and $\bar{x} = 195.5$

so $X \sim N(0, \sigma^2)$

$\bar{X} \sim N(0, \frac{\sigma^2}{10})$

$\frac{\bar{X} - 0}{\sqrt{\frac{s^2}{10}}} \sim t_9$ as we estimate σ with s_{n-1}

so p-value $= P\left(t_9 > \frac{195.5 - 0}{\sqrt{\frac{230.584^2}{10}}}\right)$

$= P(t_9 > 2.68112)$

$= 0.01258$

< 0.05

so we have evidence to reject H_0 and the mean difference of times is greater than zero. This provides evidence to support the hypothesis that alcohol reduces mean time of useful consciousness

we would alternatively perform a Wilcoxon Signed Rank Sum test on the paired data, that would mean we would not need to make the assumption about the differences being normal. We would, however, have to assume that the distributions of the two sets of data are symmetrical. This assumption may not be valid for the alcohol group due to the outlier value of 900 being present.

7. a)

Specimen	1	2	3	4	5	6	7	8
A	119	173	100	99	77	121	84	73
B	106	153	83	95	69	123	84	67
A-B	13	20	17	4	8	-2	0	6.

paired t-test.

X = difference in blood levels (where difference = analyser A − analyser B)

$X \sim N(\mu, \sigma^2)$, assumed.

$H_0: \mu = 0$

$H_1: \mu \neq 0$

Assume H_0 to be true

$\alpha = 5\%$

two tail test

we estimate σ with $s_{n-1} = 7.86947$, and we know $\bar{x} = 8.25$

so $X \sim N(0, \sigma^2)$

$\bar{X} \sim N(0, \frac{\sigma^2}{8})$

So $\frac{\bar{X} - 0}{\sqrt{\frac{s^2}{8}}} \sim t_7$ as we estimate σ with s_{n-1}

$$p\text{-value} = 2 \times P\left(t_7 > \frac{8.25 - 0}{\sqrt{\frac{7.86947^2}{8}}}\right)$$

$= 2 \times P(t_7 > 2.9652)$

$= 2 \times 0.010475$

$= 0.02095$

< 0.05

So we have evidence to reject H_0 and conclude that the mean difference in blood levels is non-zero, which suggests that the two analysers record differently.

b) i) sample of $n = 7$ had $\bar{x} = 93.5$ and $s_{n-1} = 4.7$

let X = value recorded

$X \sim N(\mu, \sigma^2)$

$H_0: \mu = 90$

$H_1: \mu \neq 90$

Assume H_0 to be true

$\alpha = 5\%$

two tailed test

we estimate σ^2 with s_{n-1} and n is small, so t-test

$X \sim N(90, \sigma^2)$

$\bar{X} \sim N(90, \frac{\sigma^2}{7})$

$\frac{\bar{X} - 90}{\sqrt{\frac{s^2}{7}}} \sim t_6$

cont/

7 cont/

$$\text{so p-value} = 2 \times P\left(t_6 > \frac{93.5-90}{\sqrt{\frac{4.7^2}{7}}}\right)$$

$$= 2 \times P(t_6 > 1.97024)$$

$$= 2 \times 0.024405$$

$$= 0.048811$$

$$< 0.05$$

So we have evidence to reject H_0 and conclude that the results have not come from a population with mean 90.

b) ii) in part (a) the machines were being compared against one another, and not for their accuracy against a controlled sample

If you are trying to establish the better analyser, the test in part (b) achieves the objective with fewer distracting factors.

If all the results were available to us, we could not only compare each machine to the standard solution of 90, but also which machine was most consistent in its measurement, with least variability.

8.

	sample size n	sum of marks	mean of marks
V	25	1060	42·4
W	15	819	54·6

let X_V = marks awarded by examiner V X_W = marks awarded by examiner W

$X_V \sim N(\mu_V, 15^2)$ $X_W \sim N(\mu_W, 15^2)$

$H_0 : \mu_V = \mu_W$

$H_1 : \mu_V < \mu_W$

Assume H_0 to be true

$\alpha = 5\%$

one tail test

we have $\bar{x}_V = 42{\cdot}5$ and $\bar{x}_W = 54{\cdot}6$

so as we know population variances we perform two sample z-test

$\bar{X}_V \sim N\left(\mu_V, \frac{15^2}{25}\right)$ $\bar{X}_W \sim N\left(\mu_W, \frac{15^2}{15}\right)$

$$\bar{X}_V - \bar{X}_W \sim N\left(\mu_V - \mu_W, \frac{15^2}{25} + \frac{15^2}{15}\right)$$

$$\text{so } \frac{\bar{X}_V - \bar{X}_W - (\mu_V - \mu_W)}{\sqrt{\frac{15^2}{25} + \frac{15^2}{15}}} \sim N(0, 1^2)$$

$$\text{so P-value} = P\left(Z < \frac{42{\cdot}5 - 54{\cdot}6 - (0)}{\sqrt{\frac{15^2}{25} + \frac{15^2}{15}}}\right)$$

$$= P(Z < -2{\cdot}4699)$$

$$= 0{\cdot}006758$$

$$< 0{\cdot}05$$

So we have evidence to reject H_0 and conclude the mean marks from examiner V are less than the mean marks from examiner W.

---

reason 1: the examiners may not have samples of scripts which were comparable – the random sample may have had bias in it.

reason 2: examiner V may have been working from marking instructions that were not as 'refined' as those used by examiner W, so they were not using the same 'rules'.

other reasons: the assumptions underlying this test may not be valid – i.e. the variances of the sample not equal, or the marks not distributed normally.

I would suggest accessing the raw data and performing a non-parametric test on the medians of the two samples, using the Mann-Whitney test.

13. a)

Property	A	B	C	D	E	F	G	H
Trainee A	83·7	58·8	77·7	85·1	91·9	66·4	69·8	48·5
Trainee B	79·6	59·2	75·8	84·3	90·1	65·2	66·9	53·8
A − B	4·1	−0·4	1·9	0·8	1·8	1·2	2·9	−5·3

let X = difference in valuations.

assume X to be normally distributed

$X \sim N(\mu, \sigma^2)$

$H_0 : \mu = 0$

$H_1 : \mu \neq 0$

Assume H_0 to be true

$\alpha = 5\%$

two tail test

we estimate σ with $s_{n-1} = 2{\cdot}83536$ and we know $\bar{x} = 0{\cdot}875$

so $X \sim N(0, \sigma^2)$

$\bar{X} \sim N(0, \frac{\sigma^2}{8})$

so $\dfrac{\bar{X} - 0}{\sqrt{\frac{s^2}{8}}} \sim t_7$ as we estimate σ with s_{n-1}

$$\text{p-value} = 2 \times P\left(t_7 > \frac{0{\cdot}875 - 0}{\sqrt{\frac{2{\cdot}83536^2}{8}}}\right)$$

$= 2 \times P(t_7 > 0{\cdot}872859)$

$= 2 \times 0{\cdot}205839$

$= 0{\cdot}411678$

$> 0{\cdot}05$

So we have no evidence to reject H_0

We do not have evidence to suggest that the mean difference in valuations is not zero.

This suggests that there is no difference in the trainee's valuations.

cont/

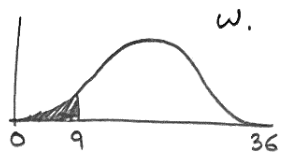
13b) we have paired data, so we perform Wilcoxon Signed Rank Sum test

H_0 : $\eta_{difference} = 0$ Assume H_0 to be true where difference = A − B.
H_1 : $\eta_{difference} \neq 0$ $\alpha = 5\%$. Two-tail test

A − B	4·1	−0·4	1·9	0·8	1·8	1·2	2·9	−5·3
|A − B|	4·1	0·4	1·9	0·8	1·8	1·2	2·9	5·3
rank	7	$\boxed{1}$	5	2	4	3	6	$\boxed{8}$

So $W_- = 1+8 = 9$
$W_+ = 7+5+2+4+3+6 = 27$ } $9+27 = 36 = \frac{1}{2} \times 8 \times 9$ ✓check

$W = \min(W_-, W_+) = 9$



we are interested in $P(W \leq 9)$

from tables, $P(W \leq 5) = 0{\cdot}05$
$P(W \leq 3) = 0{\cdot}025$
$P(W \leq 1) = 0{\cdot}01$

so $P(W \leq 9) > 0{\cdot}025$, we are not in critical region

Thus we have no evidence to reject H_0

We do not have evidence that the population median difference is not zero.

This suggests that the valuations do not differ.

15.

	with celebrity	without celebrity
n	125	75
$\sum(\text{scores})$	1275	705
$\sum(\text{scores})^2$	16485	9705.

The sample variances can be used as accurate estimates of the corresponding population variances as the sample sizes are both large. Thus instead of a t-test we can perform a z-test on the two samples.

let X_1 = scores with celebrity

$E(X_1) = \mu_1 \quad V(X_1) = \sigma_1^2$

$n_1 = 125$

$\bar{x}_1 = \frac{1275}{125} = 10.2$

$s_1^2 = \dfrac{16485 - \frac{1275^2}{125}}{124}$

$\Rightarrow s_1^2 = 28.0645 \approx \sigma_1^2$

X_2 = scores without celebrity

$E(X_2) = \mu_2 \quad V(X_2) = \sigma_2^2$

$n_2 = 75$

$\bar{x}_2 = \frac{705}{75} = 9.4$

$s_2^2 = \dfrac{9705 - \frac{705^2}{75}}{74}$

$s_2^2 = 41.5946 \approx \sigma_2^2$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 > \mu_2$

Assume H_0 to be true

$\alpha = 5\%$

by CLT, $\bar{X}_1 \approx N\left(\mu_1, \frac{\sigma_1^2}{125}\right)$ and $\bar{X}_2 \approx N\left(\mu_2, \frac{\sigma_2^2}{75}\right)$

so $\bar{X}_1 - \bar{X}_2 \approx N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{125} + \frac{\sigma_2^2}{75}\right)$

so $\dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{125} + \frac{\sigma_2^2}{75}}} \approx N(0, 1^2)$

test statistic, $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (0)}{\sqrt{\frac{s_1^2}{125} + \frac{s_2^2}{75}}}$ by approximating σ_1 with s_1, and σ_2 with s_2

$= \dfrac{(10.2 - 9.4) - 0}{\sqrt{\frac{28.0645^2}{125} + \frac{41.5946^2}{75}}}$

$= 0.14762$

p-value $= P(Z > 0.14762) = 0.441321 > 0.05$

So we do not have evidence to reject H_0

We do not have evidence that the mean scores with a celebrity are higher than those without a celebrity. This suggests that celebrity endorsement does not affect the mean persuasiveness score.