Review exercise 2

1 $\text{H}_0 : \mu = 18, \text{H}_1 : \mu < 18$

Remember to identify which is H_0 and which is H_1. This is a one-tail test since we are only interested in whether the time taken to solve the puzzle has reduced. You must use the correct parameter (μ).

$$z = \frac{(16.5 - 18)}{\left(\frac{3}{\sqrt{15}}\right)} = -1.9364\ldots$$

Using $z = \dfrac{x - \mu}{\frac{\sigma}{\sqrt{n}}}$

5% one tail c.v. is $z = -1.6449$

Use the percentage point table and quote the figure in full.

significant *or* reject H_0 *or* in critical region

$-1.9364 < -1.6449$

There is evidence that the (mean) time to complete the puzzles has reduced

or Robert is getting faster (at doing the puzzles).

State your conclusion in the context of the question.

2 **a** $\bar{x} = \dfrac{361.6}{80} = 4.52$

$$\hat{\sigma}^2 = s^2 = \frac{1753.95 - 80 \times \bar{x}^2}{79} = 1.5128$$

$$\text{or } \hat{\sigma}^2 = s^2 = \frac{80}{79} \times \left(\frac{1753.95}{80} - \bar{x}^2\right) = 1.5128$$

Using $\dfrac{\sum x^2 - n\bar{x}^2}{n-1}$

or $\dfrac{n}{n-1}\left(\dfrac{\sum x^2}{n} - \bar{x}^2\right)$

b $\text{H}_0 : \mu_A = \mu_B \quad \text{H}_1 : \mu_A > \mu_B$

This is a difference of means test. When stating hypotheses you must make it clear which mean is greater when it is a one-tailed test.

$$z = \frac{4.52 - 4.06}{\sqrt{\frac{1.5128}{80} + \frac{2.50}{60}}} = \left(\frac{0.46}{\sqrt{0.060576}}\right)$$

$= 1.8689$ or -1.8689 if $B - A$ was used.

Using $z = \dfrac{(\bar{A} - \bar{B}) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$

One tail c.v. is $z = 1.6449$

$1.87 > 1.6449$ so reject H_0.

Use the percentage point table and quote the figure in full.

There is evidence that diet A is better than diet B *or* evidence that (mean) weight lost in the first week using diet A is more than with B.

State your conclusion in the context of the question.

2 c CLT enables you to assume that $\bar{A}$ and $\bar{B}$ are normally distributed since both samples are large.

d Assumed $\sigma_A^2 = s_A^2$ and $\sigma_B^2 = s_B^2$

> Variance must be known to use the test.
>
> Remember, σ^2 is the population variance and s^2 is an unbiased estimator of the population variance.

3
$$123.5 = \bar{x} - 2.5758 \times \frac{\sigma}{\sqrt{n}} \quad \textbf{(1)}$$
$$154.7 = \bar{x} - 2.5758 \times \frac{\sigma}{\sqrt{n}} \quad \textbf{(2)}$$

> 99% confidence interval, so each tail is 0.05. Use the percentage point table and quote the figure in full. C.I.
> $\bar{x} \pm 2.5758 \times \frac{\sigma}{\sqrt{n}}$

$$\bar{x} = \frac{1}{2}(123.5 + 154.7) = 139.1$$

> Add equations **(1)** and **(2)** to find $\bar{x}$ or calculate the mean of the given limits.

$$2.5758 \times \frac{\sigma}{\sqrt{n}} = 154.7 - 139.1$$
$$= 15.6$$

> Substitute $\bar{x}$ into equation **(1)** or **(2)** to find $\frac{\sigma}{\sqrt{n}}$

$$\frac{\sigma}{\sqrt{n}} = \frac{15.6}{2.5758}$$

So 95% C.I. $= 139.1 \pm 1.9600 \times \frac{15.6}{2.5758}$

> 95% confidence interval, so each tail is 0.025.
>
> Substitute in $\bar{x}$ and $\frac{\sigma}{\sqrt{n}}$

$$= (127.22\ldots, 150.97\ldots)$$
$$= (127, 151)$$

> Answers should be given to at least 3 significant figures.

4 a
$$\bar{X} = \frac{500}{10} = 50$$
$$s^2 = \frac{25\,001.74 - 10 \times 50^2}{9}$$
$$= 0.193$$

> Using $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right)$

b For 95% confidence interval, z value is 1.96.
Confidence interval is therefore:
$$\left(50 - 1.96 \times \frac{0.5}{\sqrt{10}}, 50 + 1.96 \times \frac{0.5}{\sqrt{10}}\right)$$
$$= (49.690..., 50.309...)$$
$$= (49.7, 50.3)$$

c For 99% confidence interval, z value is 2.5758.
Confidence interval is therefore:
$$\left(50 - 2.5758 \times \frac{0.5}{\sqrt{10}}, 50 + 2.5758 \times \frac{0.5}{\sqrt{10}}\right)$$
$$= (49.592..., 50.407...)$$
$$= (49.6, 50.4)$$

5 a $\hat{\mu} = \bar{x} = \dfrac{82+98+140+110+90+125+150+130+70+110}{10}$

$= 110.5$

$s^2 = \dfrac{128153 - 10\times 110.5^2}{9}$

$= 672.28$

Using $\dfrac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\dfrac{n}{n-1}\left(\dfrac{\sum x^2}{n} - \bar{x}^2\right)$

b 95% confidence limits are:

$110.5 \pm 1.96 \times \dfrac{25}{\sqrt{10}}$

$= (95.005, 125.995)$

$= (95.0, 126)$

95% confidence interval, so each tail is 0.025. Use the percentage point table and quote the figure in full.

C.I.: $\bar{x} \pm 1.9600 \times \dfrac{\sigma}{\sqrt{n}}$

Answers should be given to at least 3 significant figures.

c $0.95^{15} = 0.46329\ldots = 0.4633$ (4 d.p.)

6 a $\bar{x} = \left(\dfrac{6046}{36} = \right) 167.94\ldots$

$s^2 = \dfrac{1016\,338 - 36 \times \bar{x}^2}{35}$

$= 27.0$

Using $\dfrac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\dfrac{n}{n-1}\left(\dfrac{\sum x^2}{n} - \bar{x}^2\right)$

b 99% confidence interval is: $\bar{x} \pm 2.5758 \times \dfrac{5.1}{\sqrt{36}}$

$= 167.94 \pm 2.5758 \times \dfrac{5.1}{\sqrt{36}}$

$= (167.75, 170.13)$

$= (166, 170)$

99% confidence interval, so each tail is 0.005. Use the percentage point table and quote the figure in full.

C.I.: $\bar{x} \pm 2.5758 \times \dfrac{\sigma}{\sqrt{n}}$

Answers should be given to at least 3 significant figures.

7 a $H_0: \mu_F = \mu_M \quad H_1: \mu_F \neq \mu_M$

This is a difference of means test.

$z = \dfrac{6.86 - 5.48}{\sqrt{\dfrac{4.51^2}{200} + \dfrac{3.62^2}{100}}}$

$= 2.860\ldots$

Using $z = \dfrac{(\bar{A} - \bar{B}) - (\mu_A - \mu_B)}{\sqrt{\dfrac{\sigma_A^2}{n_A} - \dfrac{\sigma_B^2}{n_B}}}$

2-tail 5% critical value $(\pm)$ 1.96

$2.860 > +1.96$

Use the percentage point table and quote the figure in full.

Significant result or reject the null hypothesis. There is evidence of a difference in the (mean) amount spent on junk food by male and female teenagers.

State your conclusion in the context of the question.

b CLT enables us to assume $\bar{F}$ and $\bar{M}$ are normally distributed.

8 a Let X represent repair time

$\therefore \sum x = 1435 \therefore \bar{x} = \frac{1435}{5} = 287$

$\sum x^2 = 442\,575$

$\therefore s^2 = \frac{442\,575 - 5\times 287^2}{4}$

$= 7682.5$

Using $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right)$

b $\mathrm{P}(|\mu - \hat{\mu}|) < 20 = 0.95$

$\therefore 1.96 \times \frac{\sigma}{\sqrt{n}} = 20$

$\therefore n = \frac{1.96^2\sigma^2}{20^2} = \frac{1.96^2\sigma^2}{400} = \frac{1.96^2 \times 100^2}{400} = 96.04$

$\therefore$ Sample size $(\geqslant)$ 97 required

The repair time is between 80 and 120. 95% confidence interval, so each tail is 0.025. Use the percentage point table and quote the figure in full.

C.I.: $\bar{x} \pm 1.96 \times \frac{\sigma}{\sqrt{n}}$

9 a $s^2 = \frac{2962 - 15\times\left(\frac{208}{15}\right)^2}{14} = 5.55$

$\frac{14\times 5.55}{23.685} < \sigma^2 < \frac{14\times 5.55}{6.571}$

$3.28 < \sigma^2 < 11.83$

b Since 9 lies in the interval, yes, it supports the assertion.

10 a Confidence interval $= \left(\frac{15\times 0.003}{27.488}, \frac{15\times 0.003}{6.262}\right)$

$= (0.00164, 0.00719)$

b $0.07^2 = 0.0049$

0.0049 is within the 95% confidence interval.

There is no evidence to reject the idea that the standard deviation of the volumes is not 0.07 or the machine is working well.

11 a $\mathrm{H}_0: \sigma^2 = 4; \mathrm{H}_1: \sigma^2 > 4$

$v = 19, X_{19}^2(0.05) = 30.144$

$\frac{(n-1)S^2}{\sigma^2} = \frac{19\times 6.25}{4} = 29.6875$

Since $29.6875 < 30.144$ there is insufficient evidence to reject H_0.

There is insufficient evidence to suggest that the standard deviation is greater than 2.

b That times are normally distributed.

12 $F_{10,12}(5\%) = 2.75 \therefore b = 2.75$

$$a = \frac{1}{F_{12,10}(5\%)} = \frac{1}{2.91} = 0.344$$

13
$$P(X > 2.85) = 0.05$$
$$P\left(X < \frac{1}{5.67}\right) = 0.01$$
$$\therefore P\left(\frac{1}{5.67} < X < 2.85\right) = 1 - 0.05 - 0.01$$
$$= 0.94$$

14 a $H_0 : \sigma_A^2 = \sigma_B^2, H_1 : \sigma_A^2 \neq \sigma_B^2$

critical value $F_{24,25} = 1.96$

$$\frac{s_B{}^2}{s_A{}^2} = 2.10$$

Since 2.10 is in the critical region, we reject H_0 and conclude there is evidence that the two variances are different.

b The populations of pebble lengths are normally distributed.

15 a $\left(\bar{x} = \frac{466}{4} = 116.5\right) \quad s_x^t = \frac{54386 - 4\bar{x}^2}{3} = 32.\dot{3} \text{ or } \frac{97}{3}$

$$0.216 < \frac{3s_x^2}{\sigma^2} < 9.348$$

$10.376\ldots < \sigma^2 < 449.07\ldots$ so confidence interval is (10.376, 449.07)

b $H_0 : \sigma_M^2 = \sigma_S^2 \quad H_1 : \sigma_M^2 > \sigma_S^2$

$$\frac{S_M^2}{S_s^2} = \frac{318.8}{32.\dot{3}} = 9.859....$$

$F_{6,3}(1\% \text{ c.v.}) = 27.91$

$9.859\ldots < 27.91$, so do not reject H_0. There is insufficient evidence of an increase in variance.

16 a 95% confidence interval for μ is:

$$168 \pm t_{24}(2.5\%)\sqrt{\frac{1.79}{25}} = 1.68 \pm 2.064\sqrt{\frac{1.79}{25}} = (1.13, 2.23)$$

16 b 95% confidence interval for σ^2 is:

$$12.401 < \frac{24\times 1.79}{\sigma^2} < 39.364$$

$\sigma^2 > 1.09, \sigma^2 < 3.46$

$\therefore$ confidence interval on σ^2 is (1.09, 3.46)

c Require $P(X > 2.5) = P\left(Z > \frac{2.5-\mu}{\sigma}\right)$ to be as small as possible OR

$\frac{2.5-\mu}{\sigma}$ to be as large as possible; both imply lowest σ and μ.

$$\frac{2.5-1.13}{\sqrt{1.09}} = 1.31$$

$P(Z > 1.31) = 1 - 0.9049 = 0.0951$

17 $\bar{x} = 4.01$

$s = 0.7992\ldots$

a $4.01 \pm t_9(2.5\%)\frac{0.7992..}{\sqrt{10}} = 4.01 \pm 2.262\frac{0.7992..}{\sqrt{10}}$

$= 4.5816\ldots \text{and } 3.4383\ldots$

i.e. (3.4383, 4.5816.)

b $2.700 < \frac{9\times 0.7992\ldots^2}{s^2} < 19.023$

$\sigma^2 < 2.13,\ \sigma^2 > 0.302$

i.e. $(0.302, 2.13)$

c $P(X > 7) = P\left(Z > \frac{7-\mu}{\sigma}\right)$ needs to be as high as possible.

Therefore μ and σ must be as big as possible.

Proportion with high blood glucose level $= P\left(Z > \frac{7-4.581}{\sqrt{2.13}}\right)$

$= 1 - 0.9515$

$= 0.0485$

$= 4.85\%$

18 a $\bar{x} = 123.1$

$s = 5.87745....$

(NB: $\sum x = 1231; \sum x^2 = 151\,847$)

i 95% confidence interval is given by

$$123.1 \pm 2.262 \times \frac{5.87745\ldots}{\sqrt{10}}$$

i.e. $(118.8958\ldots, 127.30418\ldots)$

18 a ii 95% confidence interval is given by:

$$\frac{9\times5.87745\ldots^2}{19.023}<\sigma^2<\frac{9\times5.87745\ldots^2}{2.700}$$

i.e. (16.34336…, 115.14806....)

b 130 is just above confidence interval.

16 is just below confidence interval.

Thus, the supervisor should be concerned about the speed of the new typist since both their average speed is too slow and the variability of the time is too large.

19 a $H_0:\sigma^2=0.9\quad H_1:\sigma^2\neq0.9$

$v=19$

CR (Lower tail 10.117)
Upper tail 30.144

Test statistic $=\frac{19\times15}{0.9}=31.6666$, significant

There is sufficient evidence that the variance of the length of spring is different from 0.9.

b $H_0:\mu=100\quad H_1:\mu>100$

$t_{19}=1.328$ is the critical value

$$t=\frac{100.6-100}{\sqrt{\frac{1.5}{20}}}=2.19$$

$2.19>1.328$ **so the result is** significant. We therefore reject H_0 and accept H_1.

The mean length of spring is greater than 100.

20 a

$$S_A^2=\frac{1}{10}(3\,960\,540)-\frac{6600^2}{11}=54.0$$

$$S_B^2=\frac{1}{12}(7\,410\,579)-\frac{9815^2}{13}=21.1\dot{6}$$

$H_0:\sigma_A^2=\sigma_B^2; H_1:\sigma_A^2\neq\sigma_B^2$

Critical region: $F_{10,12}>2.75$

$$\frac{S_A^2}{S_B^2}=\frac{54.0}{21.16}=2.55118\ldots$$

Since 2.55118... is not in the critical region we can assume that the variances are equal.

20 b $H_0 : \mu_B = \mu_A + 150; H_1 : \mu_B > \mu_A + 150$

$\text{CR} : t_{22}(0.05) > 1.717$

$$S_p^2 = \frac{10 \times 54.0 + 12 \times 21.1\dot{6}}{22} = 36.0909$$

$$t = \frac{755 - 600 - 150}{\sqrt{36.0909\ldots\left(\frac{1}{11} + \frac{1}{13}\right)}} = 2.03157$$

Since 2.03... is in the critical region, we reject H_0 and conclude that the mean weight of cauliflowers from *B* exceeds that from *A* by at least 150 g.

c Samples from normal populations

Equal variances

Independent samples

21 a i $H_0 : \sigma_c^2 = \sigma_N^2, H_1 : \sigma_C^2 > \sigma_N^2$

$$\frac{S_C^2}{s_N^2} = \frac{5.7^2}{3.5^2} = 2.652\ldots, \ F_{8,9}(5\%) \text{ critical value } = 3.23$$

$2.652\ldots < 3.23$. Not significant, so do not reject H_0.

There is insufficient evidence that variance using conventional method is greater.

ii $H_0 : \mu_N = \mu_C, H_1 : \mu_N > \mu_C$

$$s^2 = \frac{8 \times 5.7^2 + 9 \times 3.5^2}{17} = \frac{370.17}{17} = 21.774\ldots$$

$$\text{Test statistic } t = \frac{82.3 - 78.2}{\sqrt{21.774\ldots\left(\frac{1}{9} + \frac{1}{10}\right)}} = 1.912\ldots$$

$t_{17}(5\%)$ 1-tail critical value = 1.740

$1.912\ldots > 1.74$, so significant. Therefore reject H_0.

There is evidence that new style leads to an increase in the mean.

b Assumed population of marks obtained were independent and normally distributed.

c Unbiased estimate of common variance is s^2 in **ii**.

$$7.564 < \frac{17s^2}{\sigma^2} < 30.191$$

$$\sigma^2 > \frac{17 \times 21.774}{30.191} = 12.3 \text{ (1 d.p.)}$$

$$\sigma^2 < \frac{17 \times 21.774}{7.564} = 48.9 \text{ (1 d.p.)}$$

Confidence interval on σ^2 is (12.3, 48.9).

22 a $H_1: \sigma_A^2 = \sigma_B^2 \quad H_0: \sigma_A^2 \neq \sigma_B^2$

$$s_A^2 = 22.5 \quad s_B^2 = 21.6$$

$$\frac{s_A^2}{s_B^2} = 1.04$$

$$F_{(8,6)} = 4.15$$

$1.04 < 4.15$ so do not reject H_0. The variances are the same.

b Assume the samples are selected at random (independent)

c
$$s^2{}_p = \frac{8(22.5) + 6(21.62)}{14} = 22.12$$

$H_0: \mu_A = \mu_B \quad H_1: \mu_A \neq \mu_B$

$$t = \frac{40.667 - 39.57}{\sqrt{22.12}\sqrt{\frac{1}{9} + \frac{1}{7}}} = 0.462$$

Critical value = $t_{14}(2.5\%) = 2.145$

$0.462 < 2.145$, no evidence to reject H_0

The means are the same.

d Music has no effect on performance.

23 a $H_0: \sigma_R^2 = \sigma_E^2 \quad H_1: \sigma_R^2 \neq \sigma_E^2$

$F_{6,12}(5\%)_{1\,\text{tail}}\,\text{cv} = 3.00, \quad \frac{s_E^2}{s_R^2} = \frac{35.79}{14.48} = 2.4716\ldots$

Not significant, so do not reject H_0

Insufficient evidence to suspect $\sigma_R^2 \neq \sigma_E^2$

b $H_0: \mu_R = \mu_E \quad H_1: \mu_R \neq \mu_E$

$$s^2 = \frac{6 \times 35.79 \times 12 \times 14.48}{18} = 21.58\dot{3}$$

$$t = \frac{32.21 - 28.43}{s\sqrt{\frac{1}{13} + \frac{1}{7}}} = 1.78146\ldots$$

$t_{18}(5\%)_{2\,\text{tail}}\,\text{cv} = 2.101$

$\therefore$ Not significant

Insufficient evidence of difference in mean performance.

c Test in **b** requires $\sigma_1^2 \neq \sigma_2^2$

d For example: same type of driving, same roads and journey length, same weather, same driver

24 a $\bar{x} = 668.125 \quad s = 84.425$

$t_7(5\%) = 1.895$

Confidence limits $= 668.125 \pm \dfrac{1.895 \times 84.425}{\sqrt{8}} = 611.6 \text{ and } 724.7$

Confidence interval $= (612, 725)$

b Normal distribution

c £650 is within the confidence interval. No need to worry.

25 $\mathrm{H_0}: \mu = 1012 \quad \mathrm{H_1}: \mu \neq 1012$

$$\bar{x} = \frac{13\,700}{14} (= 978.57\ldots)$$

$$S_x^2 = \frac{13\,448\,750 - 14\bar{x}^2}{13} (= 3255.49)$$

$$t_{13} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{978.6 - 1012}{\frac{57.04}{\sqrt{14}}} = -2.19\ldots$$

$t_{13}(5\%)$ two-tail critical value = –2.160

Significant result – there is evidence of a change in mean weight of squirrels

26 Let x represent weight of flour.

$\sum x = 8055 \quad \therefore \bar{x} = 1006.875$

$\sum x_2 = 8\,110\,611 \therefore s^2 = \dfrac{1}{7}\left\{8\,110\,611 - \dfrac{8055^2}{8}\right\} = 33.26785...$

$\therefore s = 5.767825$

$\mathrm{H_0}: \mu = 1010; \mathrm{H_1}: \mu < 1010$

Critical value : $t = -1.895$ so critical region $t \leqslant -1.895$

$$t = \frac{(1006.875 - 1010)}{\left(\frac{5.7678}{\sqrt{8}}\right)} = -1.5324$$

Since −1.53 is not in the critical region, there is insufficient evidence to reject $\mathrm{H_0}$.

The mean weight of flour delivered by the machine is 1010 g.

27 d: 7 2 −3 1 −1 −2 10 5

$$\sum d = 19; \sum d^2 = 193$$

$$\therefore \bar{d} = \frac{19}{8} = 2.375; S_d^2 = \frac{1}{7}(193 - \frac{19^2}{8}) = 21.125$$

$$H_0 : \mu_D = 0; H_1 : \mu_D > 0$$

$$t = \frac{2.375 - 0}{\sqrt{\frac{21.125}{8}}} = 1.4615\ldots$$

$$v = 7 \Rightarrow \text{critical region} : t \geq 1.895$$

Since 1.4915... is *not* in the critical region, there is insufficient evidence to reject H_0 and we conclude that there is insufficient evidence to support the doctors' belief.

28 a d = coursework – written: 4, –3, –3, 4, 6, 3, –4, 17, 7, 7

$$\bar{d} = \frac{38}{10} = 3.8, s_d^2 = \frac{498 - 10\bar{d}^2}{9} = 39.2\dot{8}$$

$$\text{test statistic} : t = \frac{3.8}{\frac{s_d}{\sqrt{10}}} = 1.917\ldots$$

$$H_0 : \mu_d = 0 \quad H_1 : \mu_d > 0$$

$t_9(5\%)$ c.v. is 1.833;

$\therefore$ significant – there is evidence coursework marks are higher

b The difference between the marks follows a normal distribution.

29 $D = \text{dry} - \text{wet} \qquad H_0 : \mu_D = 0, H_1 : \mu_D \neq 0$

$d : 0.6, -1, -1.9, -1.4, -1.3, 0.5, -1.6, -0.6, -1.8$

$$\bar{d} : -\frac{8.5}{9} = -0.9\dot{4},\ s_d^2 = \frac{15.03 - 9 \times \left(\bar{d}\right)^2}{8} = 0.87527\ldots$$

$$t = \frac{-0.9\dot{4}}{\frac{s_d}{\sqrt{9}}} = \text{awrt} -3.03$$

t_8 2-tail 1% critical value = 3.355

Not significant – insufficient evidence of a difference between mean strength.

30 a $H_0 : \mu_d = 0$

$H_1 : \mu_d > 0$

where d = without solar heating – with solar heating

$d = 6 \quad -3 \quad 7 \quad -2 \quad -8 \quad 6 \quad 5 \quad 11 \quad 5$

$\bar{d} = 3$

$s_d = 6$

$n_d = 9$

$$\therefore \text{ test statistic} = \frac{(3-0)}{\left(\frac{6}{\sqrt{9}}\right)}$$

t.s. = 1.5

Critical value = $t_8(5\%) = 1.860$

So critical region: $t > 1.860$

Test statistic not in critical region, so accept H_0. Conclude there is insufficient evidence that solar heating reduces mean weekly fuel consumption.

b The differences are normally distributed.

31 a

$$s_p^2 = \frac{7\times 7.84 + 7\times 4}{7+7} = 5.92$$

$$s_p = 2.433105$$

$$H_0 : \mu_A = \mu_B, H_1 : \mu_A \neq \mu_B$$

$$t = \frac{26.125 - 25}{2.43\sqrt{\frac{1}{8}+\frac{1}{8}}} = 0.92474$$

$t_{14}(2.5\%) = 2.145$

Insufficient evidence to reject H_0.

Conclude that there is no difference in the means.

b $d = 2, 5, -2, 1, 3, -4, 1, 3$

$$\bar{d} = \frac{9}{8} = 1.125$$

$$s_d^2 = \frac{69 - 8\times 1.125^2}{7} = 8.410714$$

$$H_0 : \delta = 0, H_1 : \delta \neq 0$$

$$t = \frac{1.125}{\sqrt{\frac{8.41}{8}}} = 1.0972$$

$t_7(2.5\%) = 2.365$

There is no significant evidence of a difference between method A and method B.

c Paired sample, as they are two measurements on the same orange.

32 a Confidence interval is given by:

$$\bar{x} \pm t_{19} \times \frac{s}{\sqrt{n}}$$

i.e. $207.1 \pm 2.539 \times \sqrt{\frac{3.2}{20}}$

i.e. 207.1 ± 1.0156

i.e. $(206.08\ldots, 208.1156)$

b $\bar{x}_G = \frac{2046.2}{10} = 204.62$

$$S_p^2 = \frac{19 \times 3.2 + 9 \times 10.217\dot{3}}{28}$$

$$= 5.45557\ldots$$

Confidence interval is given by:

$$\bar{x}_B - \bar{x}_G \pm t_{28} \times \sqrt{5.45557\left(\frac{1}{20} + \frac{1}{10}\right)}$$

i.e. $(207.1 - 204.62) \pm 2.154\sqrt{5.45557\left(\frac{1}{20} + \frac{1}{10}\right)}$

i.e. 2.48 ± 1.94854

i.e. $(0.53146, 4.42854)$

Challenge

1 a $\mathrm{E}\left(\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3\right) = \frac{2}{3}\mu - \frac{1}{2}\mu + \frac{5}{6}\mu = \mu$

$\mathrm{E}(Y) = \mu \Rightarrow$ unbiased

b $\mathrm{E}(aX_1 + bX_2) = a\mu + b\mu = \mu$

$a + b = 1$

$$\mathrm{Var}(aX_1 + bX_2) = a^2\sigma^2 + b^2\sigma^2$$

$$= a^2\sigma^2 + (1-a)^2\sigma^2$$

$$= (2a^2 - 2a + 1)\sigma^2$$

c Minimum value when $(4a - 2)\ \sigma^2 = 0$ (from differentiation)

$\Rightarrow 4a - 2 = 0$

$$a = \frac{1}{2}, b = \frac{1}{2}$$

Challenge

2 a Consider the cumulative distribution functions:

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leqslant x \leqslant 1 \\ 1 & x > 1 \end{cases}$$

In order for $M < x$ you must have no more than 1 of the $X_i > x$ so either all three $X_i < x$ with probability $[F(x)]^3 = x^3$
or exactly one of the $X_i > x$ with probability $3 \times [F(x)]^2 \times (1 - F(x)) = 3x^2(1 - x)$
So $P(M < x) = x^3 + 3x^2(1 - x) = 3x^2 - 2x^3$
So M has cumulative distribution function

$$H(x) = \begin{cases} 0 & x < 0 \\ 3x^2 - 2x^3 & 0 \leqslant x \leqslant 1 \\ 1 & x > 1 \end{cases}$$

Differentiating to find h(x):

$$h(x) = \begin{cases} 6x - 6x^2 & 0 \leqslant x \leqslant 1 \\ 0 & \text{otherwise} \end{cases}$$

b $E(M) = \int_0^1 x(6x - 6x^2)dx = \left[2x^3 - \frac{3}{2}x^4\right]_0^1 = 0.5 = E(X)$

c $E(M^2) = \int_0^1 x^2(6x - 6x^2)dx = \left[\frac{3}{2}x^4 - \frac{6}{5}x^5\right]_0^1 = 0.3$

$$\begin{aligned} \text{Var}(M) &= E\left(M^2\right) - \left(E\left(M\right)\right)^2 \\ &= 0.3 - 0.5^2 \\ &= 0.05 \end{aligned}$$

Hence standard error of $M = \sqrt{0.05} = 0.224$ (3 s.f.)