Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 1

Question:

Find the upper 5% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 12, b = 18$,

b $a = 4, b = 11$,

c $a = 6, b = 9$.

Solution:

a 2.34
b 3.36
c 3.37

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise A, Question 2

Question:

Find the lower 5% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 6, b = 8$,

b $a = 25, b = 12$,

c $a = 5, b = 5$.

Solution:

a $\frac{1}{F_{8,6}} = 0.241$

b $\frac{1}{F_{12,25}} = 0.463$

c $\frac{1}{F_{5,5}} = 0.198$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 3

Question:

Find the upper 1% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 12, b = 18$,

b $a = 6, b = 16$,

c $a = 5, b = 9$.

Solution:

a 3.37

b 4.20

c 6.06

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 4

Question:

Find the lower 1% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 3, b = 12$,

b $a = 8, b = 12$,

c $a = 5, b = 12$.

Solution:

a $\frac{1}{F_{12,3}} = 0.0370$

b $\frac{1}{F_{12,8}} = 0.176$

c $\frac{1}{F_{12,5}} = 0.101$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 5

Question:

Find the lower and upper 5% critical value for a $F_{a,b}$-distribution in each of the following cases:

a $a = 8, b = 10$,

b $a = 12, b = 10$,

c $a = 3, b = 5$.

Solution:

a $3.07, \frac{1}{F_{10,8}} = 0.299$

b $2.91, \frac{1}{F_{10,12}} = 0.364$

c $5.41, \frac{1}{F_{5,3}} 0.111$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 6

Question:

The random variable X follows a $F_{40,12}$-distribution.
Find $\mathrm{P}(X < 0.5)$.

Solution:

$$\begin{aligned}\mathrm{P}(X < 0.5) &= \mathrm{P}(F_{40,12} < 0.5)\\ &= \mathrm{P}(F_{12,40} > \frac{1}{0.5})\\ &= \mathrm{P}(F_{12,40} > 2)\end{aligned}$$

From the tables $F_{12,40}(0.05) = 2$

$\therefore \mathrm{P}(F_{12,40} > 2) = \mathrm{P}(F_{40,12} < 0.5) = 0.05$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 7

Question:

The random variable X follows a $F_{12,8}$-distribution.

Find $\mathrm{P}\left(\frac{1}{2.85} < X < 3.28\right)$

Solution:

$$\begin{aligned}
\mathrm{P}(X < 3.28) &= 1 - \mathrm{P}(F_{12,8} > 3.28) \\
&= 1 - 0.05 = 0.95
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(X < \frac{1}{2.85}) &= \mathrm{P}(F_{12,8} < \frac{1}{2.85}) \\
&= \mathrm{P}(F_{8,12} > 2.85)
\end{aligned}$$

$$\therefore \mathrm{P}(X < \frac{1}{2.85}) = 0.05$$

$$\begin{aligned}
\mathrm{P}\left(\frac{1}{2.85} < X < 3.28\right) &= \mathrm{P}(X < 3.28) - \mathrm{P}\left(X < \frac{1}{2.85}\right) \\
&= 0.95 - 0.05 \\
&= 0.90
\end{aligned}$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 8

Question:

The random variable X has an F-distribution with 2 and 7 degrees of freedom.
Find $P(X < 9.55)$. **[E]**

Solution:

$$\begin{aligned} P(X < 9.55) &= 1 - P(F_{2,7} > 9.55) \\ &= 1 - 0.01 \\ &= 0.99 \end{aligned}$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise A, Question 9

Question:

The random variable X follows an F-distribution with 6 and 12 degrees of freedom.

a Show that $P(0.25 < X < 3.00) = 0.9$.

A large number of values are randomly selected from an F-distribution with 6 and 12 degrees of freedom.

b Find the probability that the seventh value to be selected will be the third value to lie between 0.25 and 3.00. **[E]**

Solution:

a
$$\begin{aligned} P(X < 3.00) &= 1 - P(F_{6,12} > 3.00) \\ &= 1 - 0.05 = 0.95 \\ P(X > 0.25) &= P(F_{12,6} > \frac{1}{0.25}) \\ &= P(F_{12,6} > 4) \\ &= 0.95 \\ P(0.25 < X < 3) &= 0.95 - 0.05 \\ &= 0.90 \end{aligned}$$

b $^6C_2(0.9)^2(0.1)^4 \times 0.9 = 0.00109$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 1

Question:

Random samples are taken from two normally distributed populations. There are 11 observations from the first population and the best estimate for the population variance is $s^2 = 7.6$. There are 7 observations from the second population and the best estimate for the population variance is $s^2 = 6.4$.

Test, at the 5% level of significance, the hypothesis $\mathrm{H}_0 : \sigma_1{}^2 = \sigma_2{}^2$ against the alternative hypothesis $\mathrm{H}_1 : \sigma_1{}^2 > \sigma_2{}^2$.

Solution:

Critical value is $F_{10,6} = 4.06$

$F_{\text{test}} = \frac{7.6}{6.4} = 1.1875$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_1^2 = \sigma_2^2$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 2

Question:

Random samples are taken from two normally distributed populations. There are 25 observations from the first population and the best estimate for the population variance is $s^2 = 0.42$. There are 41 observations from the second population and the best estimate for the population variance is $s^2 = 0.17$.
Test, at the 1% significance level, the hypothesis $\mathrm{H}_0 : {\sigma_1}^2 = {\sigma_2}^2$ against the alternative hypothesis $\mathrm{H}_1 : {\sigma_1}^2 > {\sigma_2}^2$.

Solution:

Critical value is $F_{24,40} = 2.29$

$F_{\text{test}} = \dfrac{0.42}{0.17} = 2.4706$

In critical region

reject H_0 – there is evidence to suggest that $\sigma_1^2 > \sigma_2^2$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 3

Question:

The variance of the lengths of a sample of 9 tent-poles produced by a machine was 63 mm^2. A second machine produced a sample of 13 tent-poles with a variance of 225 mm^2. Both these values are unbiased estimates of the population variances.

a Test, at the 10 % level, whether there is evidence that the machines differ in variability, stating the null and alternative hypotheses.

b State the assumption you have made about the distribution of the populations in order to carry out the test in part **a**. **[E]**

Solution:

a $H_0: \sigma_1^2 = \sigma_2^2 \quad H_1: \sigma_1^2 \neq \sigma_2^2$

Critical value is $F_{12,8} = 3.28$

$F_{\text{test}} = \frac{225}{63} = 3.57$

In critical region

reject H_0 – There is evidence to suggest that the machines differ in variability

b Population distributions are assumed to be normal

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 4

Question:

Random samples are taken from two normally distributed populations. The size of the sample from the first population is $n_1 = 13$ and this gives an unbiased estimate for the population variance $s_1^2 = 36.4$. The figures for the second population are $n_2 = 9$ and $s_2^2 = 52.6$.

Test, at the 5% significance level, whether $\sigma_1^2 = \sigma_2^2$ or if $\sigma_1^2 > \sigma_2^2$.

Solution:

Critical value is $F_{8,12} = 2.85$

$F_{\text{test}} = \dfrac{52.6}{36.4} = 1.445$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_1^2 = \sigma_2^2$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 5

Question:

Dining chairs Ltd are in the process of selecting a make of glue for using on the joints of their furniture. There are two possible contenders – Goodstick which is the more expensive, and Holdtight, the cheaper of the two.

The company are concerned that, while both glues are said to have the same adhesive power, one might be more variable than the other.

A series of trials are carried out with each glue and the joints tested to destruction. The force in newtons at which each joint failed is recorded. The results are as follows:

Goodstick	10.3	8.2	9.5	9.9	11.4	
Holdtight	9.6	10.8	9.9	10.8	10.0	10.2

a Test, at the 10% significance level, whether or not the variances are equal.
b Which glue would you recommend and why?

Solution:

a $\sigma^2_{\text{goodstick}} = 1.363$

$\sigma^2_{\text{Holdtight}} = 0.24167$

Critical value is $F_{4,5} = 5.19$

$F_{\text{test}} = \frac{1.363}{0.24167} = 5.64$

In critical region

reject H_0 – there is evidence to suggest that the variances are not equal.

b Holdtight as it is less variable and cheaper.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise B, Question 6

Question:

The closing balances, £x, of a number of randomly chosen bank current accounts of two different types, Chegrit and Dicabalk, are analysed by a statistician. The summary statistics are given in the table below.

	Sample size	$\sum x$	$\sum x^2$
Chegrit	7	276	143 742
Dicaballk	15	394	102 341

Stating clearly your hypotheses test, at the 10% significance level, whether or not the two distributions have the same variance. (You may assume that the closing balances of each type of account are normally distributed.) **[E]**

Solution:

$\sigma^2_{\text{Chegrit}} = 22\,143.286$

$\sigma^2_{\text{Dicabalk}} = 6570.85238$

Critical value is $F_{6,14} = 2.85$

$$F_{\text{test}} = \frac{22143.286}{6570.85238} = 3.3699$$

In critical region − there is evidence to suggest that their variances differ

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise B, Question 7

Question:

Bigborough council wishes to change the bulbs in their traffic lights at regular intervals so that there is a very small probability that any light bulb will fail in service. The council are anxious that the length of time between changes should be as long as possible, and to this end they have obtained a sample of bulbs from another manufacturer, who claims the same bulb life as their present manufacturer. The council wishes therefore to select the manufacturer whose bulbs have the smallest variance.

When they last tested a random sample of 9 bulbs from their present supplier the summary results were $\sum x = 9415$ hours, $\sum x^2 = 9\,863\,681$, where x represents the lifetime of a bulb.

A random sample of 8 bulbs from the prospective new supplier gave the following bulb lifetimes in hours: 1002, 1018, 943, 1030, 984, 963, 1048, 994.

a Calculate unbiased estimates for the means and variances of the two populations.

Assuming that the lifetimes of bulbs are normally distributed,

b test, at the 10% significance level, whether or not the two variances are equal.

c State your recommendation to the council, giving reasons for your choice.

Solution:

a $\mu_1 = 1046 \quad s_1^2 = 1818.11$ and $\mu_2 = 997.75 \quad s_2^2 = 1200.21$

b Critical value is $F_{8,7} = 3.73$

$$F_{\text{test}} = \frac{1818.111}{1200.21} = 1.5148$$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_1^2 = \sigma_2^2$

c Use present supplier who appears to have a higher mean.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise C, Question 1

Question:

A random sample of 10 toothed winkles was taken from a sheltered shore, and a sample of 15 was taken from a non-sheltered shore. The maximum basal width, (x mm), of the shells was measured and the results are summarised below.

Sheltered shore: $\bar{x} = 25, s^2 = 4$. Non-sheltered shore: $\bar{x} = 22, s^2 = 5.3$.

a Find a 95% confidence interval for the difference between the means.

b State an assumption that you have made when calculating this interval.

Solution:

a

$$s_p^2 = \frac{(9 \times 4) + (14 \times 5.3)}{10 + 15 - 2} = 4.7913$$

$$S_p = 2.189$$

$$t_{23}(2.5\%) = 2.069$$

$$(25 - 22) \pm 2.069 \times 2.189\sqrt{\frac{1}{15} + \frac{1}{10}} = 3 \pm 1.849$$

$$= (1.151, 4.849)$$

b Independent random samples, normal distributions, common variance

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 2

Question:

A packet of plant seeds was sown and, when the seeds had germinated and begun to grow, 8 were transferred into pots containing a soil-less compost and 10 were grown on in a soil-based compost. After 6 weeks of growth the heights, x, in cm of the plants were measured with the following results:
Soil-less compost: 9.3, 8.7, 7.8, 10.0, 9.2, 9.5, 7.9, 8.9.
Soil-based compost: 12.8, 13.1, 11.2, 10.1, 13.1, 12.0, 12.5, 11.7, 11.9, 12.0.
Assuming that the populations are normally distributed, and that there is a difference between the two means calculate a 90% confidence interval for this difference.

Solution:

$$\bar{x}_s = 8.9125, \quad s_s^2 = 0.58125$$

$$\bar{x}_{ns} = 12.04 \quad s_{ns}^2 = 0.84933$$

$$s_p^2 = \frac{(7\times 0.58125)+(9\times 0.84933)}{10+8-2} = 0.7319$$

$$s_p = 0.855$$

$$t_{16}(5\%) = 1.746$$

$$(12.04-8.9125)\pm 1.746\times 0.855\sqrt{\frac{1}{10}+\frac{1}{8}} = 3.1275\pm 0.7081$$

$$= (2.419, 3.836)$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 3

Question:

Forty children were randomly selected from all 12-year-old children in a large city to compare two methods of teaching the spelling of 50 words which were likely to be unfamiliar to the children. Twenty children were randomly allocated to each method. Six weeks later the children were tested to see how many of the words they could spell correctly. The summary statistics for the two methods are given in the table below, where $\bar{x}$ is the mean number of words spelt correctly, s^2 is an unbiased estimate of the variance of the number of words spelt correctly and n is the number of children taught using each method.

	$\bar{x}$	s^2	n
Method A	32.7	6.1^2	20
Method B	38.2	5.2^2	20

a Calculate a 99% confidence interval for the difference between the mean numbers of words spelt correctly by children who used Method B and Method A.
b State two assumptions you have made in carrying out part **a**.
c Interpret your result.

Solution:

a

$$s_p^2 = \frac{(19\times 6.1^2)+(19\times 5.2^2)}{20+20-2} = 32.125$$

$$s_p = 5.66789$$

$$t_{38}(0.5\%) = 2.712$$

$$(38.2-32.7)\pm 2.712\times 5.66789\sqrt{\frac{1}{20}+\frac{1}{20}} = 5.5\pm 4.86083$$

$$= (0.6392, 10.3608)$$

b normality and equal variances
c zero not in interval $\Rightarrow$ method B seems better than method A

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise C, Question 4

Question:

The table below shows summary statistics for the mean daily consumption of cigarettes by a random sample of 10 smokers before and after their attendence at an anti-smoking workshop with $\bar{x}$ representing the means and s^2 representing the unbiased estimates of population variance in each case.

	$\bar{x}$	s^2	n
Mean daily consumption before the workshop	18.6	32.488	10
Mean daily consumption after the workshop	14.3	33.344	10

Stating clearly any assumption you make, calculate a 90% confidence interval for the difference in the mean daily consumption of cigarettes before and after the workshop.

Solution:

Assume same variances and that the population of differences is normally distributed

$$s_p^2 = \frac{(9\times 32.488)+(9\times 33.344)}{10+10-2} = 32.916$$

$$s_p = 5.73725$$

$$t_{18}(5\%) = 1.734$$

$$(18.6-14.3)\pm 1.734\times 5.73725\sqrt{\frac{1}{10}+\frac{1}{10}} = 4.3\pm 4.44905$$

$$= (-0.149, 8.749)$$

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 1

Question:

A random sample of size 20 from a normal population gave $\bar{x} = 16, s^2 = 12$.

A second random sample of size 11 from a normal population gave $\bar{x} = 14, s^2 = 12$.

a Assuming that the both populations have the same variance, find an unbiased estimate for that variance.

b Test, at the 5% level of significance, the suggestion that the two populations have the same mean.

Solution:

a $s_p^2 = \dfrac{(19\times12)+(10\times12)}{20+11-2} = 12 \quad \text{so } s_p = \sqrt{12} = 3.464$

b $H_0: \mu_{1st} = \mu_{2nd} \quad H_1: \mu_{1st} \neq \mu_{2nd}$

critical value $t_{29}(0.025) = 2.045$

critical region is $t \le -2.045$ and $t \ge 2.045$

$$t = \frac{(16-14)-0}{3.464\sqrt{\frac{1}{20}+\frac{1}{11}}} = 1.538$$

Not in critical region – do not reject H_0

There is evidence to suggest that the populations have the same mean

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 2

Question:

Salmon reared in Scottish fish farms are generally larger than wild salmon.
A fisherman measured the length of the first 6 salmon caught on his boat at a fish farm. Their lengths in centimetres were

42.8, 40.0, 38.2, 37.5, 37.0, 36.5.

Chefs prefer wild salmon to fish-farmed salmon because of their better flavour. A chef was offered 4 salmon that were claimed to be wild. Their lengths in centimetres were

42.0, 43.0, 41.5, 40.0.

Use the information given above and a suitable t-test at the 5% level of significance to help the chef to decide if the claim is likely to be correct. You may assume that the populations are normally distributed.

Solution:

$\text{H}_0 : \mu_c = \mu_F \quad \text{H}_1 : \mu_c > \mu_F$

$\bar{x}_F = 38.67, \quad s_F{}^2 = 5.5827$

$\bar{x}_c = 41.625 \quad s_c{}^2 = 1.5625$

$$s_p{}^2 = \frac{(5 \times 5.5827) + (3 \times 1.5625)}{6 + 4 - 2} = 4.075 \quad \text{so } s_p = \sqrt{4.075} = 2.0187$$

critical value $t_8(0.05) = 1.86$

critical region $t \geq 1.860$

$$t = \frac{(41.625 - 38.67) - 0}{2.0187\sqrt{\frac{1}{6} + \frac{1}{4}}} = 2.270$$

In the critical region − reject H_0 there is evidence to suggest that the salmon are wild

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 3

Question:

In order to check the effectiveness of three drugs against the E. *coli* bacillus, 15 cultures of the bacillus (5 for each of 3 different antibiotics) had discs soaked in the antibiotics placed in their centre. The 15 cultures were left for a time and the area in cm^2 per microgram of drug where the E. *coli* was killed was measured. The results for three different drugs are given below:

Streptomycin	0.210, 0.252, 0.251, 0.210, 0.256, 0.253
Tetracycline	0.123, 0.090, 0.123, 0.141, 0.142, 0.092
Erythromycin	0.134, 0.120, 0.123, 0.210, 0.134, 0.134

a It was thought that Tetracycline and Erythromycin seemed equally as effective. Assuming that the populations are normally distributed, test this at the 5% significance level.

b Streptomycin was thought to be more effective than either of the others. Treating the other 2 as being a single sample of 12, test this assertion at the same level of significance.

Solution:

a $\bar{x}_t = 0.1185$, $s_t^2 = 0.0005227$.

$\bar{x}_e = 0.1425$ $s_e^2 = 0.0011319$.

$$s_p^2 = \frac{(5\times 0.0005227)+(5\times 0.0011319)}{6+6-2} = 0.000827 \text{ so } s_p = \sqrt{0.000827} = 0.02876$$

$H_0: \mu_t = \mu_e$ $H_1: \mu_t \neq \mu_e$

critical value $t_{10}(0.025) = 2.228$

critical region $t \le -2.228$ and $t \ge 2.228$

$$t = \frac{(0.1425 - 0.1185) - 0}{0.02876\sqrt{\frac{1}{6}+\frac{1}{6}}} = 1.445$$

not in the critical region − accept H_0; there is evidence to suggest that Tetracycline and Erythromycin are equally as effective

b $\bar{x}_s = 0.2387$, $s_s^2 = 0.0004959$

$\bar{x}_2 = 0.1305$, $s_2^2 = 0.000909$

$$s_p^2 = \frac{(11\times 0.000909)+(5\times 0.0004959)}{12+6-2} = 0.000780, \text{ so } s_p = \sqrt{0.000780} = 0.0279$$

$H_0: \mu_s = \mu_2$ $H_1: \mu_s > \mu_2$

critical value $t_{16}(0.05) = 1.746$

critical region $t \ge 1.746$

$$t = \frac{(0.2387 - 0.1305) - 0}{0.0279\sqrt{\frac{1}{12}+\frac{1}{6}}} = 7.75$$

in the critical region − reject H_0. There is evidence to suggest that Streptomycin is more effective than the others.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise D, Question 4

Question:

To test whether a new version of a computer programming language enabled faster task completion, the same task was performed by 16 programmers, divided at random into two groups. The first group used the new version of the language, and the time for task completion, in hours, for each programmer was as follows:

4.9 6.3 9.6 5.2 4.1 7.2 4.0.

The second group used the old version, and their times were summarised as follows:

$n = 9, \sum x = 71.2, \sum x^2 = 604.92$.

a State the null and alternative hypotheses.

b Perform an appropriate test at the 5% level of significance.

In order to compare like with like, experiments such as this are often performed using the same individuals in the first and the second groups.

c Give a reason why this strategy would not be appropriate in this case. **[E]**

Solution:

a $H_0 : \mu_{\text{old}} = \mu_{\text{new}}, \quad H_1 : \mu_{\text{old}} > \mu_{\text{new}}$

b $\bar{x}_{old} = 7.911 \quad s_{old}^2 = 5.206$

$\bar{x}_{new} = 5.9, \quad s_{new}^2 = 3.98$

$$s_p{}^2 = \frac{(6 \times 3.98) + (8 \times 5.206)}{7 + 9 - 2} = 4.6806 \quad \text{so } s_p = \sqrt{4.6806} = 2.1635$$

Critical value $t_{14} = 1.761$

critical region $t \geq 1.761$

$$\text{Test statistic } t = \frac{(7.911 - 5.9) - 0}{2.1635\sqrt{\frac{1}{9} + \frac{1}{7}}} = 1.84446$$

Significant – there is evidence to suggest that new language does improve time.

c Once task is solved the programmer should be quicker next time with either language.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise D, Question 5

Question:

A company undertakes investigations to compare the fuel consumption, x, in miles per gallon, of two different cars, the Volcera and the Spintono, with a view to purchasing a number as company cars.
For a random sample of 12 Volceras the fuel consumption is summarised by
$\sum V = 384$ and $\sum V^2 = 12\,480$.
A statistician incorrectly combines the figures for the sample of 12 Volceras with those of a random sample of 15 Spintonos, then carries out calculations as if they are all one larger sample and obtains the results $\bar{y} = 34$ and $s^2 = 23$.

a Show that, for the sample of 15 Spintonos, $\sum x = 534$ and $\sum x^2 = 19\,330$.

Given that the variance of the fuel consumption for each make of car is σ^2

b obtain an unbiased estimate for σ^2.

c Test, at the 5% level of signifance, whether there is a difference between the mean fuel consumption of the two models of car. State your hypothese and conclusion clearly.

d State any further assumption you made in order to be able to carry out your test in c.

e Give two precautions which could be taken when undertaking an investigation into the fuel consumption of two models of car to ensure that a fair comparison is made.

[E]

Solution:

a $27 \times 34 - 384 = 534 = \sum x$

$$23 = \frac{\sum y^2}{27-1} - \frac{918^2}{27(27-1)}$$

$384 + 534 = 918$

$\sum y^2 = 31\,810$

$31\,810 - 12\,480 = 19\,330 = \sum x^2$

b $\bar{x}_v = 32 \quad s_v^2 = 17.45$

$\bar{x}_s = 35.6, \quad s_s^2 = 22.829$

$$s_p^{\,2} = \frac{(14 \times 22.829) + (11 \times 17.45)}{12 + 15 - 2} = 20.464$$

c $H_0 : \mu_v = \mu_s, \quad H_1 : \mu_v \neq \mu_s$

Critical value $t_{25}(0.025) = 2.060$

Critical region $t \leq -2.060$ and $t \leq 2.060$

$$t = \frac{(35.6 - 32) - 0}{4.524\sqrt{\frac{1}{15} + \frac{1}{12}}} = 2.0547$$ — accept H_0 — no evidence to suggest difference in means

d normality

e same types of driving, roads and weather

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 1

Question:

It is claimed that completion of a shorthand course has increased the shorthand speeds of the students.

a If the suggestion that the mean speed of the students has not altered is to be tested, write down suitable hypotheses for which **i** a two-tailed test is appropriate, and **ii** a one-tailed test is appropriate.

The table below gives the shorthand speeds of students before and after the course.

Student	A	B	C	D	E	F
Speed before in words/minute	35	40	28	45	30	32
Speed after	42	45	28	45	40	40

b Carry out a paired t-test, at the 5% significance level, to determine whether or not there has been an increase in shorthand speeds.

Solution:

a **i** $H_0 : \mu_D = 0, H_1 : \mu_D \neq 0$

ii $H_0 : \mu_D = 0, H_1 : \mu_D > 0$

b $\sum d = 30 \quad \sum d^2 = 238$

$\bar{d} = 5$

$s^2 = \frac{238 - 6(5)^2}{5} = 17.6$

$s = 4.195$

Critical value $t_5(5\%) = 2.015$

The critical region is $t > 2.015$

$t = \frac{5-0}{\frac{4.195}{\sqrt{6}}}$

$= 2.919$

In the critical region – reject H_0.

There is evidence to suggest that there has been an increase in shorthand speed.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 2

Question:

A large number of students took two General Studies papers that were supposed to be of equal difficulty. The results for 10 students chosen at random are shown below:

Candidate	A	B	C	D	E	F	G	H	I	J
Paper 1	18	25	40	10	38	20	25	35	18	43
Paper 2	20	27	39	12	40	23	20	35	20	41

The teacher looked at the marks of a random sample of 10 students, and decided that paper 2 was easier than paper 1.
Given that the marks on each paper are normally distributed, carry out an appropriate test, at the 1% level of significance.

Solution:

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

$\sum d = 5 \quad \sum d^2 = 59$

$\bar{d} = 0.5$

$s^2 = \frac{59 - 10(0.5)^2}{9} = 6.278$

$s = 2.50555$

Critical value $t_9(1\%) = 2.821$

The critical region is $t > 2.821$.

$t = \frac{0.5 - 0}{\frac{2.50555}{\sqrt{10}}}$

$= 0.631$

Not in the critical region. Do not reject H_0.

There is insufficient evidence to suggest that paper 2 is easier than paper 1 so the teacher is not correct.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 3

Question:

It is claimed by the manufacturer that by chewing a special flavoured chewing gum smokers are able to reduce their craving for cigarettes, and thus cut down on the number of cigarettes smoked per day. In a trial of the gum on a random selection of 10 people the no-gum smoking rate and the smoking rate when chewing the gum were investigated, with the following results:

Person	A	B	C	D	E	F	G	H	I	J
Without gum smoking rate cigs./day	20	35	40	32	45	15	22	30	34	40
With gum smoking rate cigs./day	15	25	35	30	45	15	14	25	28	34

a Use a paired t-test at the 5% significance level to test the manufacturer's claim.
b State any assumptions you have had to make.

Solution:

a $H_0 : \mu_D = 0, H_1 : \mu_D > 0$

$\sum d = 47 \quad \sum d^2 = 315$

$\bar{d} = 4.7$

$$s^2 = \frac{315 - 10(4.7)^2}{9} = 10.456$$

$s = 3.234$

Critical value $t_9(5\%) = 1.833$

The critical region is $t > 1.833$

$$t = \frac{4.7 - 0}{\frac{3.234}{\sqrt{10}}}$$

$= 4.596$

In the critical region. Reject H_0.

There is evidence to suggest that chewing the gum does reduce the craving for cigarettes.

b The differences are normally distributed

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 4

Question:

The council of Somewhere town are going to put a new traffic management scheme into operation in the hope that it will make travel to work in the mornings quicker for most people. Before the scheme is put into operation, 10 randomly selected workers are asked to record the time it takes them to come into work on a Wednesday morning. After the scheme is put into place, the same 10 workers are again asked to record the time it takes them to come into work on a particular Wednesday morning.
The times in minutes are shown in the table below:

Worker	A	B	C	D	E	F	G	H	I	J
Before	23	37	53	42	39	60	54	85	46	38
After	18	35	49	42	34	48	52	79	37	37

Test, at the 5% significance level, whether or not the journey time to work has decreased.

Solution:

$H_0 : \mu_D = 0, \quad H_1 : \mu_D > 0$

$\sum d = 46 \quad \sum d^2 = 336$

$\bar{d} = 4.6$

$s^2 = \frac{336 - 10(4.6)^2}{9} = 13.8222$

$s = 3.7178$

Critical value $t_9(5\%) = 1.833$

The critical region is $t > 1.833$

$t = \frac{4.6 - 0}{\frac{3.7178}{\sqrt{10}}}$

$= 3.913$

In the critical region. Reject H_0.

There is evidence to suggest that the journey times have decreased.

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 5

Question:

A teacher is anxious to test the idea that students' results in mock examinations are good predictors for their results in actual examinations. He selects 8 students at random from those doing a mock Statistics examination and records their marks out of 100; later he collects the same students' marks in the actual examination. The resulting marks are as follows:

Student	A	B	C	D	E	F	G	H
Mock examination mark	35	86	70	91	45	64	78	38
Actual examination	45	77	81	86	53	71	68	46

a Use a paired t-test to investigate whether or not the mock examination is a good predictor. (Use a 10% significance level.)

b State any assumptions you have made.

Solution:

a $H_0 : \mu_D = 0, H_1 : \mu_D \neq 0$

$\sum d = 20 \quad \sum d^2 = 604$

$\bar{d} = 2.5$

$$s^2 = \frac{604 - 8(2.5)^2}{7} = 79.1429$$

$s = 8.896$

Critical value $t_7(5\%) = 1.895$

The critical regions are $t < -1.895$ and $t > 1.895$.

$$t = \frac{2.5 - 0}{\frac{8.896}{\sqrt{8}}}$$

$= 0.795$

Not in the critical region. Do not reject H_0.

The mock examination is a good predictor.

b The differences are normally distributed

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise E, Question 6

Question:

The manager of a dress-making company took a random sample of 10 of his employees and recorded the number of dresses made by each. He discovered that the number of dresses made between 3.00 and 5.00 p.m. was fewer than the same employees achieved between 9.00 and 11.00 a.m. He wondered if a tea break from 2.45–3.00 p.m. would increase productivity during these last two hours of the day. The number of dresses made by these workers in the last two hours of the day before and after the introduction of the tea break were as shown below.

Worker	A	B	C	D	E	F	G	H	I	J
Before	75	73	75	81	74	73	77	75	75	72
After	80	84	79	84	85	84	78	78	80	83

a Why was the comparison made for the same ten workers?

b Conduct, at the 5% level of significance, a paired t-test to see if the introduction of a tea break has increased production between 3.00 and 5.00 p.m.

Solution:

a Different people will have different productivity rates. Need a common link if want to compare before and after. This reduces experimental error due to differences between individuals so that, if a difference does exist, it is more likely to be detected.

b $H_0 : \mu_D = 0, H_1 : \mu_D > 0$

$\sum d = 65 \quad \sum d^2 = 569$

$\bar{d} = 6.5$

$$s^2 = \frac{569 - 10(6.5)^2}{9} = 16.278$$

$s = 4.0346$

Critical value $t_9(5\%) = 1.833$

The critical region is $t > 1.833$

$$t = \frac{6.5 - 0}{\frac{4.0346}{\sqrt{10}}}$$

$= 5.095$

In the critical region. Reject H_0.

There is evidence to suggest a tea break increases the number of garments made.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise E, Question 7

Question:

A drug administered in tablet form to help people sleep and a placebo was given for two weeks to a random sample of eight patients in a clinic. The drug and the placebo were given in random order for one week each. The average numbers of hours sleep that each patient had per night with the drug and with the placebo are given in the table below.

Patient	1	2	3	4	5	6	7	8
Hours of sleep with drug	10.5	6.7	8.9	6.7	9.2	10.9	11.9	7.6
Hours of sleep with placebo	10.3	6.5	9.0	5.3	8.7	7.5	9.3	7.2

Test, at the 1% level of significance, whether or not the drug increases the mean number of hours sleep per night. State your hypotheses clearly. **[E]**

Solution:

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

$\sum d = 8.6 \sum d^2 = 20.78$

$\bar{d} = 1.075$

$$s^2 = \frac{20.78 - 8(1.075)^2}{7} = 1.64786$$

$s = 1.2837$

Critical value $t_7(1\%) = 2.998$

The critical region is $t > 2.998$

$$t = \frac{1.075 - 0}{\frac{1.2837}{\sqrt{8}}}$$

$= 2.3686$

Not in the critical region. Do not reject H_0.

There is no evidence to suggest that the drug increases the mean number of hours sleep per night.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 1

Question:

The random variable X has an F-distribution with 5 and 10 degrees of freedom. Find values of a and b such that $\mathrm{P}(a \leq X \leq b) = 0.90$ **[E]**

Solution:

$\mathrm{P}(F_{5,10} \geq 3.33) = 0.05 \Rightarrow b = 3.33$

$\mathrm{P}\left(F_{10,5} \geq 4.74\right) = 0.05 \Rightarrow \mathrm{P}\left(F_{5,10} \leq \frac{1}{4.74}\right) = 0.05$

$\therefore a = 0.2110 \text{ (4 s.f.)}$

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 2

Question:

A chemist has developed a fuel additive and claims that it reduces the fuel consumption of cars. To test this claim, 8 randomly selected cars were each filled with 20 litres of fuel and driven around a race circuit. Each car was tested twice, once with the additive and once without. The distance, in miles, that each car travelled before running out of fuel are given in the table below.

Car	1	2	3	4	5	6	7	8
Distance without additive	163	172	195	170	183	185	161	176
Distance with additive	168	185	187	172	180	189	172	175

Assuming that the distances travelled follow a normal distribution and stating your hypotheses clearly test, at the 10% level of significance, whether or not there is evidence to support the chemist's claim. **[E]**

Solution:

$d: 5, 13, -8, 2, -3, 4, 11, -1$

$(\Sigma d = 23, \Sigma d^2 = 409) \quad \bar{d} = 2.875, \; sd = 6.9987 (\approx 7.00)$

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

$$t = \frac{(2.875 - 0)}{\frac{6.9987}{\sqrt{8}}} = 1.1618\ldots (\approx 1.16)$$

Critical value $t_7(10\%) = 1.415$ (1 tail)

Critical region is $t > 1.415$

Not significant

Insufficient evidence to support the chemist's claim

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 3

Question:

The standard deviation of the length of a random sample of 8 fence posts produced by a timber yard was 8 mm. A second timber yard produced a random sample of 13 fence posts with a standard deviation of 14 mm.

a Test, at the 10% significance level, whether or not there is evidence that the lengths of fence posts produced by these timber yards differ in variability. State your hypotheses clearly.

b State an assumption you have made in order to carry out the test in part **a**. **[E]**

Solution:

a $\text{H}_0 : \sigma_1^2 = \sigma_2^2$; $\text{H}_1 : \sigma_1^2 \neq \sigma_2^2$

$$\frac{s_1^2}{s_2^2} = \frac{14^2}{8^2} = 3.0625 \left(\text{or } \frac{8^2}{14^2} = 0.32653\ldots \right)$$

Critical value $F_{12,7} = 3.57$ $\left(\text{Critical value} : F_{7,12} = \frac{1}{3.57} = 0.28011 \right)$

Since 3.0625 is not in the critical region there is insufficient evidence to reject H_0.
There is insufficient evidence of a difference in the variances of the lengths of the fence posts.

b The distribution of the population of lengths of fence posts is normally distributed.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 4

Question:

A farmer set up a trial to assess the effect of two different diets on the increase in the weight of his lambs. He randomly selected 20 lambs. Ten of the lambs were given diet *A* and the other 10 lambs were given diet *B*. The gain in weight, in kg, of each lamb over the period of the trial was recorded.

a State why a paired t-test is not suitable for use with these data.

b Suggest an alternative method for selecting the sample which would make the use of a paired t-test valid.

c Suggest two other factors that the farmer might consider when selecting the sample.

The following paired data were collected.

Diet *A*	5	6	7	4.6	6.1	5.7	6.2	7.4	5	3
Diet *B*	7	7.2	8	6.4	5.1	7.9	8.2	6.2	6.1	5.8

d Using a paired t-test at the 5% significance level, test whether or not there is evidence of a difference in the weight gained by the lambs using diet *A* compared with those using diet *B*.

e State, giving a reason, which diet you would recommend the farmer to use for his lambs. **[E]**

Solution:

a The data were not collected in pairs.

b Use data from twin lambs.

c Age, weight, gender

d $d = B - A$

$d: 2, 1.2, 1, 1.8, -1, 2.2, 2, -1.2, 1.1, 2.8$

$\Sigma d = 11.9; \quad \Sigma d^2 = 30.01$

$\therefore \bar{d} = 1.19; s^2 = 1.761 \quad (s = 1.327)$

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

$$t = \frac{1.19 - 0}{\sqrt{\frac{1.761}{10}}} = 2.83574\ldots$$

$u = 9$; critical value: $t = 2.262$

Since $2.8357\ldots$ is in the critical region ($t > 2.262$) there is evidence to reject H_0.

The (mean) weight gained by the lambs is different for each diet.

e Diet B - it has the higher mean

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 5

Question:

A medical student is investigating two methods of taking a person's blood pressure. He takes a random sample of 10 people and measures their blood pressure using an arm cuff and a finger monitor. The table below shows the blood pressure for each person, measured by each method.

Person	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Arm cuff	140	110	138	127	142	112	122	128	132	160
Finger monitor	154	112	156	152	142	104	126	132	144	180

a Use a paired t-test to determine, at the 10% level of significance, whether or not there is a difference in the mean blood pressure measured using the two methods. State your hypotheses clearly.

b State an assumption about the underlying distribution of measured blood pressure required for this test. **[E]**

Solution:

a d : 14 2 18 25 0 −8 4 4 12 20

$\left(\sum d = 91, \quad \sum x^2 = 1789\right)$

$\bar{d} = 9.1 \quad s = \sqrt{106.7} = 10.332..$

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

$$t = \frac{(9.1 - 0)}{\frac{10.332}{\sqrt{10}}} = 2.785$$

Critical value $t_9 = \pm 1.833$

critical regions: $t \leq -1.833$ or $t \geq 1.833$

Significant. There is a difference between *blood pressure* measured by arm cuff and finger monitor.

b The *difference in measurements* of blood pressure is *normally distributed*

Solutionbank S4

Edexcel AS and A Level Modular Mathematics

Exercise F, Question 6

Question:

The lengths, x mm, of the forewings of a random sample of male and female adult butterflies are measured. The following statistics are obtained from the data.

	Number of butterflies	Sample mean $\bar{x}$	$\sum x^2$
Females	7	50.6	17 956.5
Males	10	53.2	28 335.1

a Assuming the lengths of the forewings are normally distributed, test, at the 10% level of significance, whether or not the variances of the two distributions are the same. State your hypotheses clearly.

b Stating your hypotheses clearly test, at the 5% level of significance, whether the mean length of the forewings of the female butterflies is less than the mean length of the forewings of the male butterflies. **[E]**

Solution:

a $H_0: \sigma^2_F = \sigma^2_M \quad H_1: \sigma^2_F \neq \sigma^2_M$

$$s^2_F = \frac{1}{6}(17\,956.5 - 7\times 50.6^2) = \frac{33.98}{6} = 5.66333\ldots$$

$$s^2_M = \frac{1}{9}(28\,335.1 - 10\times 53.2^2) = \frac{32.7}{9} = 3.63333\ldots$$

$$\frac{s^2_F}{s^2_M} = 1.5587\ldots(\text{Reciprocal } 0.6415)$$

$$F_{6,9} = 3.37(\text{or } 0.297)$$

Not in critical region. There is no reason to doubt the *variances* of the two distributions *are the same*

b $H_0: \mu_F = \mu_M \quad H_1: \mu_F < \mu_M$

$$\text{Pooled estimate } s^2 = \frac{6\times 5.66333\ldots + 9\times 3.63333}{15}$$

$$= 4.44533$$

$$s = 2.11$$

$$t = \frac{50.6 - 53.2}{2.11\sqrt{\frac{1}{7}+\frac{1}{10}}} = -2.50$$

Critical value $t_{15}(5\%) = -1.753$

so critical region $t \leq -1.753$

Significant. The mean length of the *female's forewing is* less than the length of the male's forewing

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 7

Question:

The weights, in grams, of mice are normally distributed. A biologist takes a random sample of 10 mice. She weighs each mouse and records its weight.
The ten mice are then fed on a special diet. They are weighted again after two weeks.
Their weights in grams are as follows:

Mouse	A	B	C	D	E	F	G	H	I	J
Weight before diet	50.0	48.3	47.5	54.0	38.9	42.7	50.1	46.8	40.3	41.2
Weight after diet	52.1	47.6	50.1	52.3	42.2	44.3	51.8	48.0	41.9	43.6

Stating your hypotheses clearly, and using a 1% level of significance, test whether or not the diet causes an increase in the mean weight of the mice. **[E]**

Solution:

Differences 2.1 −0.7 2.6 −1.7 3.3 1.6 1.7 1.2 1.6 2.4

$\sum d = 14.1 \quad \sum d^2 = 40.65 \quad \bar{d} = 1.41$

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

$$s = \sqrt{\frac{40.65 - 10 \times 1.41^2}{9}} = 1.5191\ldots$$

$$t = \frac{1.41}{\left(\frac{1.519\ldots}{\sqrt{10}}\right)} = 2.935$$

$t_9(1\%) = 2.821$

so critical region $t > 2.821$

2.935.. > 2.821 Evidence to reject H_0.

There has been an increase in the mean weight of the mice.

Solutionbank S4
Edexcel AS and A Level Modular Mathematics

Exercise F, Question 8

Question:

A hospital department installed a new, more sophisticated piece of equipment to replace an ageing one in the hope that it would speed up the treatment of patients. The treatment times of random samples of patients during the last week of operation of the old equipment and during the first week of operation of the new equipment were recorded. The summary results, in minutes, were:

	n	$\sum x$	$\sum x^2$
Old equipment	10	225	5136.3
New equipment	9	234	6200.0

a Show that the values of s^2 for the old and new equipment are 8.2 and 14.5 respectively.

Stating clearly your hypotheses, test

b whether the variance of the times using the new equipment is greater than the variance of the times using the old equipment, using a 5% significance level,

c whether there is a difference between the mean times for treatment using the new equipment and old equipment, using a 2% significance level.

d Find 95% confidence limits for the mean difference in treatment times between the new and old equipment.

Even if the new equipment would eventually lead to a reduction in treatment times, it might be that to begin with treatment times using the new equipment would be higher than those using the old equipment.

e Give one reason why this might be so.

f Suggest how the comparison between the old and new equipment could be improved. **[E]**

Solution:

a $s_o^2 = \frac{5136.3}{9} - \frac{(225)^2}{10(10-1)} = 8.2$

$s_n^2 = \frac{6200}{8} - \frac{(234)^2}{9(9-1)} = 14.5$

b $H_0 : \sigma_o^2 = \sigma_n^2 \quad H_1 : \sigma_o^2 < \sigma_n^2$

Critical value is $F_{8,9} = 3.23$

so critical region, $F \geq 3.23$

$F_{test} = \frac{14.5}{8.2} = 1.768$

not in critical region

accept H_0 – there is evidence to suggest that $\sigma_o^2 = \sigma_n^2$

c $s_p^2 = \frac{(9 \times 8.2) + (8 \times 14.5)}{10 + 9 - 2} = 11.1647$

$H_0 : \mu_0 = \mu_n \quad H_1 : \mu_o \neq \mu_n$

critical value $t_{17}(0.01) = 2.567$

critical region $t \leq -2.567$ and $t \geq 2.567$

$$t = \frac{(26 - 22.5) - 0}{\sqrt{11.1647}\sqrt{\frac{1}{10} + \frac{1}{9}}} = 2.2798$$

Not in the critical region – do not reject H_0.

There is evidence to suggest that there is no difference in mean times between the old and new equipment.

d $t_{17}(2.5\%) = 2.110$

$$(26 - 22.5) \pm 2.110 \times \sqrt{11.1647} \times \sqrt{\frac{1}{10} + \frac{1}{9}} = 3.5 \pm 3.2394$$

$$= (0.261, 6.739)$$

e Need to learn how to use new equipment efficiently

f Gather data on new equipment after it has been mastered