STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 1

(a)

Pig Number	1	2	3	4	5	6	7	8	9	10	11
B- A Weight difference(x)	+ 1.5	+ 0.5	+ 3.8	+ 5.2	+ 3.7	+ 1.0	- 6.0	- 4.5	- 3.5	+ 1.1	+ 0.3
x^2	2.25	0.25	14.44	27.04	13.69	1.0	36.0	20.25	12.25	1.21	0.09

1 mark for correct differences (must show negatives);
1 mark for correct squared values;   **2**

(b) (i) $\bar{x}^2 = 0.0794$; (add B-A, divide by 11, then square)
$\Sigma x^2 = 128.47$; (allow 128.5)   **2**

(ii) $s^2 = \dfrac{128.47}{11} - 0.0794 = 11.6$; $s = 3.406$; (allow 3.41)   **2**

(iii) $t = \dfrac{0.28\sqrt{11-1}}{3.406}$; $= 0.26$;   **2**

(iv) accept the null hypothesis;
calculated value is less than critical value;   **2**

(c) (i) anaemia is a lack of haemoglobin/red blood cells;
thus ability of blood to transport oxygen/carbon dioxide is impaired;
thus less available energy/ATP for protein synthesis/cellular respiration may be slowed up;   **2**
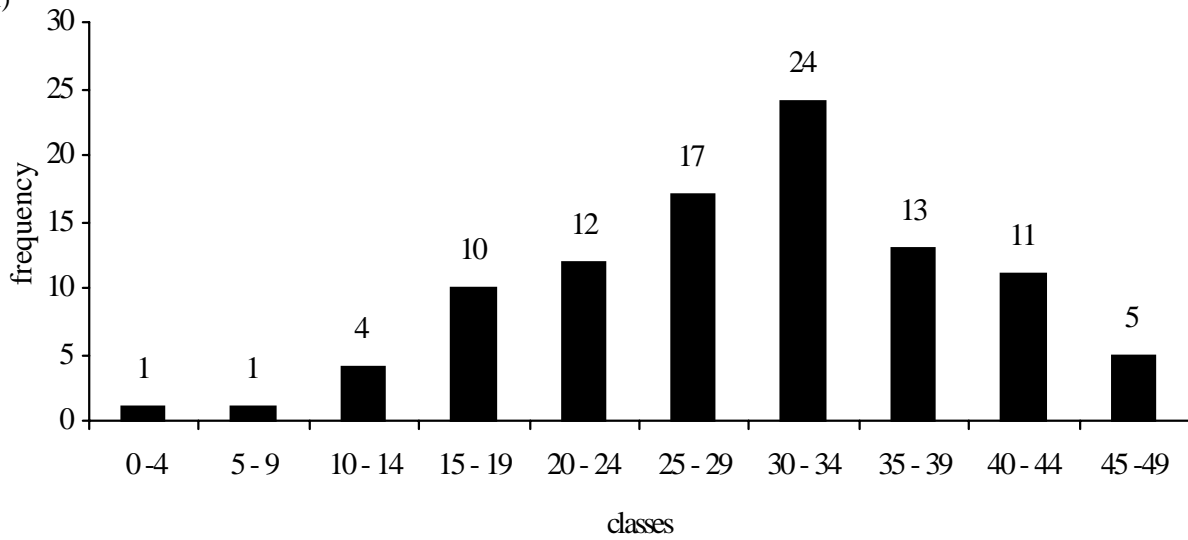
(ii) by injection (since dietary supplement has no significant effect);   **1**

TOTAL 13

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 2

(a) (i)

Class (5 lbs)	No of pigs
0 - 4	1
5 - 9	1
10 - 14	4
15 - 19	10
20 - 24	12
25 - 29	17
30 - 34	24
35 - 39	13
40 - 44	11
45 - 49	5

correct numbers of pigs written in table;
axes correctly labelled including class details;
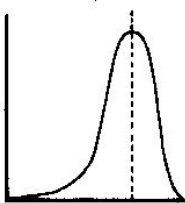suitable scale;
accurate plotting;
blocks drawn clearly/tidily;

5

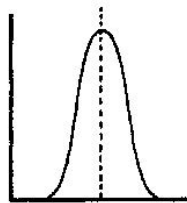(ii)   normal/top hat distribution/slightly skewed to the right; **1**
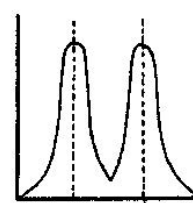
(iii) $\dfrac{2890}{98}$;   = 29.49 ; **2**

(b)

Directional	Stabilising	Disruptive
(accept other direction)	(must be narrower than orginal)	

3 correctly drawn curves;;;   3 means;;;   **6**

TOTAL 14

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 3

(a)(i)

	Correct	Incorrect
Nose closed	139.5 ;	60.5 ;
Nose open	139.5 ;	60.5 ;

4

(ii) $\chi^2 = \dfrac{(167 - 139.5)^2}{139.5} + \dfrac{(33 - 60.5)^2}{60.5} + \dfrac{(112 - 139.5)^2}{139.5} + \dfrac{(88 - 60.5)^2}{60.5}$

$= 5.421 + 12.50 + 5.421 + 12.50$; (allow mark here or in correct substitution above)

$\chi^2 = 35.84$; **2**

(iii) $n = 1$; **1**

(iv) accept the hypothesis/effect of nasal breathing is important;
since calculated value of Chi^2 is much higher than critical value; **2**

(b) (i) mouths should be rinsed out with water before each test;
cinnamon infusion should be of same concentration throughout test/concentrated enough to taste;
test subjects should not have colds/nasal congestion/be the same age/sex;
test subjects should not eat for a few hours prior to the test (so that residual tastes do not interfere with the test);**max 3**

(ii) many substances to be tasted give off volatile chemicals;
which can depolarise/be sensed by/stimulate the olfactory/smell receptors in the nasal mucosa/lining;
thus adding to the effect of the chemical depolarising/stimulating the taste buds on the tongue; **3**

TOTAL 15

QUESTIONSHEET 4

(a) (i)

	Green	White
Number of plants		
O	49	18
E	50.25;	16.75;

2

(ii) $\chi^2 = \dfrac{(49 - 50.25)^2}{50.25} + \dfrac{(18 - 16.75)^2}{16.75}$;

$= 0.124$;
(allow the working mark at any correct stage in the calculation) **2**

(iii) $n = 1$; **1**

(iv) reject the hypothesis/deviation from ratio is not significant;
calculated value of χ^2 is less than critical value (at 0.05 probability level); **2**

(b) suitable symbols, eg. G for green, g for white;

P   Gg   x   Gg   (no mark)

gametes   Ⓖ ⓖ   Ⓖ ⓖ;

F$_1$ GG   Gg   Gg   gg ;

        3 green   1 white ; **4**

TOTAL 11

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 5

(a)(i)

plant	gold accumulated/mg gold Kg^{-1} dry mass		$x = b - a$	x^2
	no thiocyanate(a)	**with thiocyanate(b)**		
1	1.01	0.99	-0.02	0.0004
2	1.09	1.09	0	0.0000
3	0.98	1.25	0.27	0.0729
4	0.71	1.34	0.63	0.3969
5	1.15	1.22	0.07	0.0049
6	1.21	1.18	-0.03	0.0009
7	1.18	1.23	0.05	0.0025
8	0.89	1.34	0.45	0.2025
9	1.26	1.21	-0.05	0.0025
10	1.07	0.97	-0.1	0.0100
11	1.30	1.18	-0.12	0.0144
12	1.17	1.23	0.06	0.0036

x values;
$\sum x = 1.21$;
$\bar{x} = \dfrac{1.21}{12} = 0.101$;   **3**

(ii) x^2 values;
$\sum x^2 = 1.464$;   **2**

(iii) $(\bar{x})^2 = 0.010(2)$;   **1**

(iv) $s^2 = \dfrac{1.464}{12} - 0.010 = 0.112$;

$s = 0.335$;   **2**

(v) $t = \dfrac{0.101\sqrt{12-1}}{0.335}$; $= 1.00$; (allow 0.999)   **2**

(b) (i) accept the null hypothesis;
calculated value is less than the critical value (at 0.05 probability level/at all probability levels shown);   **2**

(ii) use same strain/seed batch/similar Brassica seedlings;
mine waste material should have identical gold contents to begin with;
use a standard concentration of thiocyanate;
grow under similar conditions of temperature/humidity/water availability;   **max 2**

TOTAL 14

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 6

(a) the colours of the seeds produced are not in the ratio 5:4:4:2 ;   **1**

(b)

Colour of flower	Red	White	Yellow	Pink
Working	$\dfrac{5 \times 60}{15}$	$\dfrac{4 \times 60}{15}$	$\dfrac{4 \times 60}{15}$	$\dfrac{2 \times 60}{15}$
Expected numbers	20	16	16	8

   correct working ;
   correct answers;   **2**

(c) $\chi^2 = \dfrac{(15-20)^2}{20} + \dfrac{(20-16)^2}{16} + \dfrac{(10-16)^2}{16} + \dfrac{(15-8)^2}{8}$; (allow for any correct stage of working)

   $\chi^2 = 10.625$;   **2**

(d) 3:   **1**

(e) accept the null hypothesis/colours are not distributed in stated ratio;
   calculated value of χ^2 is more than the critical value;   **2**

(f) heterozygous;
   blending inheritance/incomplete dominance;
   allele for red interacts with allele for white (producing pink);   **max 2**

   TOTAL 10

QUESTIONSHEET 7

(a) 70% of dandruff sufferers will be able to clear up their dandruff with one application of the shampoo;   **1**

(b) 200 – (6.156 + 26.46 + 72.03 + 33.614) ;
   = 61.74 ;   **2**

(c) $\chi^2 = \dfrac{(15-6.156)^2}{6.156} + \dfrac{(24-26.46)^2}{26.46} + \dfrac{(57-61.74)^2}{61.74} + \dfrac{(75-72.03)^2}{72.03} + \dfrac{(29-33.614)^2}{33.614}$;
     (allow working mark at any correct stage of calculation)

   = (12.706 + 0.229 + 0.364 + 0.122 + 0.633) = 14.054 ;   **2**

(d) 4; (column 0 has to be amalgamated with column 1)   **1**

(e) accept the hypothesis/the manufacturers claim is correct;
   since calculated value of χ^2 is higher than the critical value;   **2**

   TOTAL 8

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 8

(a) the mean height of year 8 pupils at the teacher's school is less than 1.53 metres;   **1**

(b) (i)

Pupil Number	1	2	3	4	5	6	7	8	9	10
Height difference/m	- 0.03	+ 0.04	- 0.04	- 0.07	0	+ 0.09	0	- 0.02	+ 0.04	- 0.09

       1 mark for + or - signs; 1 mark for figures;   **2**

(ii) $\sum x = -0.08$; $\bar{x} = -0.008$; $\bar{x}^2 = 0.000064$; $\sum x^2 = 0.0272$;   **4**

(iii) $s^2 = \dfrac{0.0272}{10} - 0.000064 = 0.002656$;
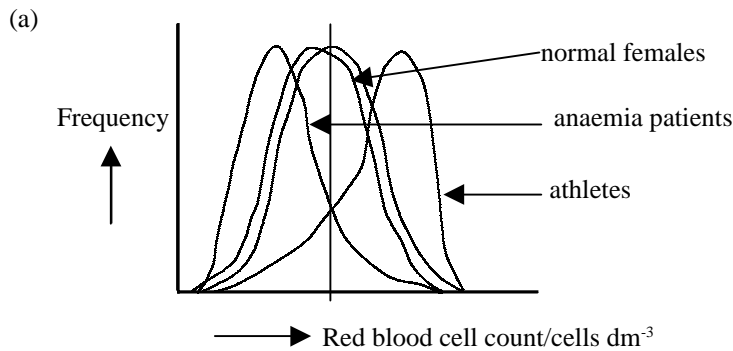
    $s = 0.0515363$;   **2**

(iv) $t = \dfrac{0.008\sqrt{10-1}}{0.0515363}$; $= 0.466$;   **2**

(v) 9;   **1**

(c) reject the hypothesis/the teacher is not correct;
calculated value is less than the critical value;   **2**

TOTAL 14

STATISTICS FOR BIOLOGY

ANSWERS & MARK SCHEMES
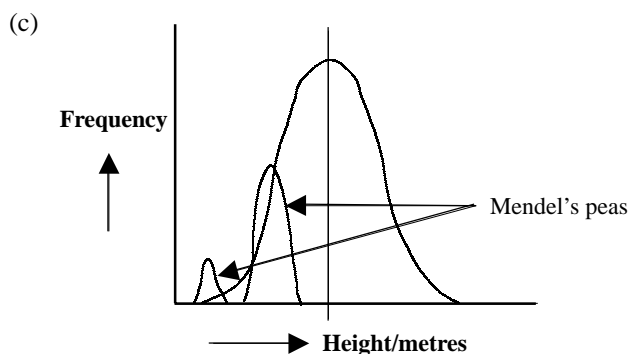
QUESTIONSHEET 9

(a)

Red blood cell count/cells dm^{-3}

- (i) correct graph;
  women have lower red cell counts than men (so normal distribution lies to the left of men); **2**

- (ii) correct graph;
  anaemic patients have fewer red cells than normal (so distribution skewed to the left);
  due to defeciency of iron/folic acid/vit B$_{12}$/cyanocobalamin/erythropoietin/due to kidney disease; **3**

- (iii) correct graph;
  red cell count has increased due to low oxygen tension (so distribution skewed to the right);
  ref to low oxygen tension stimulating erythropoietin release which stimulates red cell formation/mitosis in red bone marrow; **3**

(b)

Life span/years
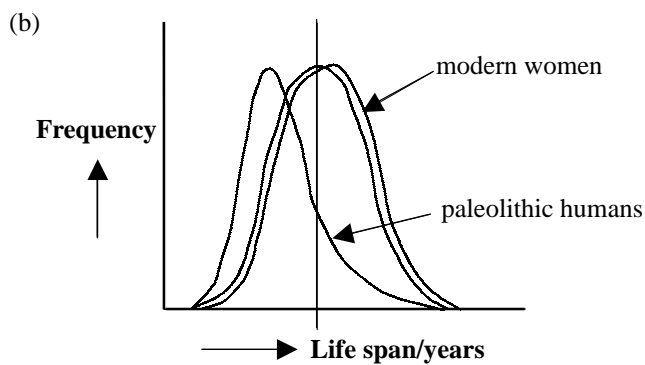
- (i) correct graph;
  women tend to live longer than men (so normal distribution lies to the right of men); **2**

- (ii) correct graph;
  paleolithic humans tended to die young/by the age of 25 (accept any reasonable value);
  due to lack of food/no health care/no medical care/killed by wild animals/any reasonable suggestion; **3**

(c)

Height/metres

correct graph;
shows discontinuous variation/bimodal distribution;
tall and short peas in the ratio of three to one/ref to monohybrid inheritance; **3**

TOTAL 16

STATISTICS FOR BIOLOGY

ANSWERS & MARK SCHEMES

QUESTIONSHEET 10

(a) (i)   the inherited characteristics do not behave in a Mendelian manner;   **1**

    (ii)

Phenotype	Observed	Expected	$(O - E)$	$(O - E)^2$
Grey winged	180	145;	35	1225
Black vestigial	52	16;	36	1296
Grey vestigial	14	48;	-34	1156
Black wild	12	48;	-36	1296

    $(O - E)$ all correct;

    $(O - E)^2$ all correct;   **6**

  (iii) $\chi^2 = \dfrac{1225}{145} + \dfrac{1296}{16} + \dfrac{1156}{48} + \dfrac{1296}{48}$;

        $= (8.45 + 81 + 24.08 + 27) = 140.53$;   **2**
    (allow the working mark at any correct stage in the calculation)

  (iv) 3;   **1**

  (v)   reject the null hypothesis/results are not a Mendelian ratio;
       calculated value is more than the critical value;   **2**

(b) the characteristics are linked;
    grey to winged and black to vestigial wings;
    alleles on the same (homologous pair of)chromosomes;
    ref to recombinants produced by chiasma formation/crossover;   **max 3**

           TOTAL 15

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 11

(a) dandelion and ribwort do not tend to grow in association with each other in the meadow; **1**

(b) (i)

		Dandelion		
		present	absent	row total
Ribwort	present	O 88 E **86.8;**	O 52 E **53.2;**	140
	absent	O 36 E **37.2;**	O 24 E **22.8;**	60
	column total	124	76	200 (grand total)

4

(ii) $\chi^2 = \dfrac{(88-86.8)^2}{86.8} + \dfrac{(52-53.2)^2}{53.2} + \dfrac{(36-37.2)^2}{37.2} + \dfrac{(24-22.8)^2}{22.8}$;

  $= (0.0166 + 0.0271 + 0.0387 + 0.0632)$

  $= 0.1456$; **2**

(iii) 1; **1**

(iv) accept the null hypothesis/there is no association between the two plants;
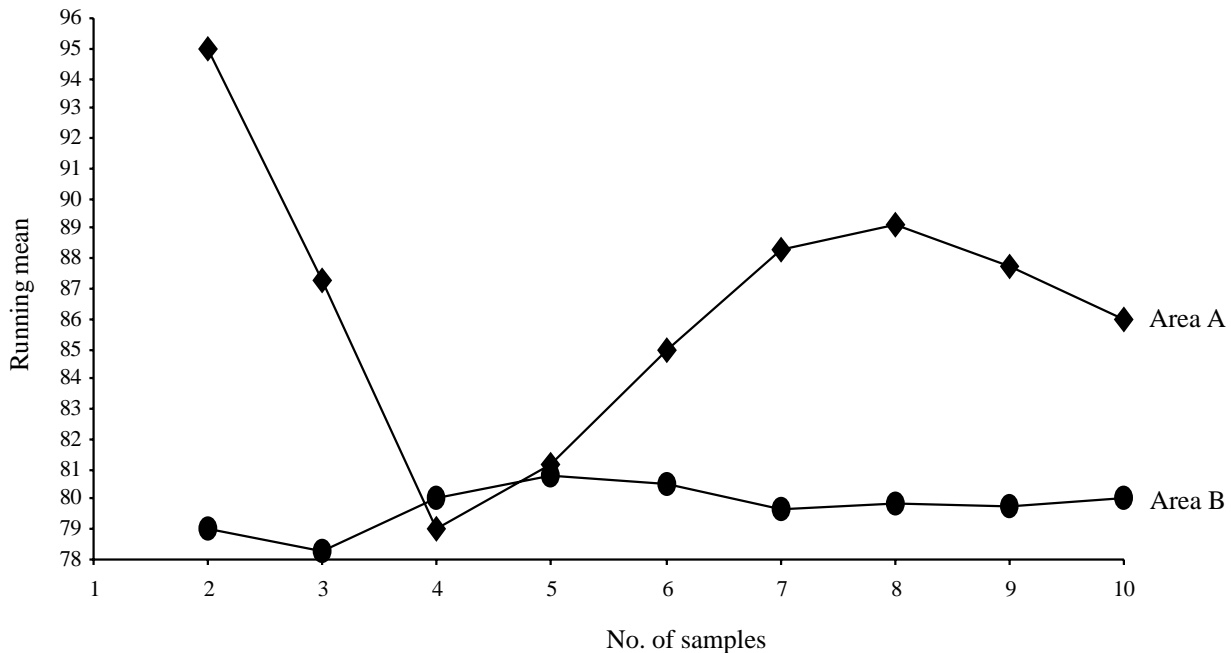   calculated value is smaller than the critical value; **2**

TOTAL 10

STATISTICS FOR BIOLOGY
ANSWERS & MARK SCHEMES

QUESTIONSHEET 12

(a) (i)   correct labelled axes;
        suitable scale;
        accurate plotting;
        points joined with a straight line; (IOB recommendations)
        curves labelled;   **5**

(ii)   area A;
       the data is much more variable in A (than in B);
       thus A would need many samples to get an accurate estimation of the mean/only a few samples in B would give an accurate
       estimation of the mean;   **3**

(iii) the time available for the investigation;
       the time required to obtain each sample;
       the ease with which each sample can be collected/many small plants more difficult to count than a few large plants;
       how much damage/disturbance the sampling may cause/trampling/crushing of delicate/rare plants;   **max 2**

(b) (i) Chi^2 ;
        used to test 'goodness of fit' of genetical data to Mendelian/linkage/expected ratios;   **2**

(ii)   t-test;
       used to test the significance of differences in the means of two (or more) sets of data;   **2**

(iii) Chi^2 ;
       used to test association between factors/light intensity and bluebell distribution;   **2**

TOTAL 16

(Current syllabuses only specify t-test and Chi^2 test for possible exam use)