Mathematical Biology Exam Paper 2021

Answers and solutions

Note: For queries about these solutions, or for solutions to any questions not included here, please contact the relevant lecturer or lecturers.

MB 2021 A1 Solution

(i) A study reports that the range of body temperatures, measured in degrees Celsius, for individuals infected with disease X is normally distributed with mean $\mu = 38.5$ and standard deviation $\sigma = 0.65$. What is the probability that a randomly selected infected individual has a temperature between 37.9°C and 39.8°C?

Answer:

$X \;\; N(\mu = 38.5, \sigma = 0.65)$

Compute z-score $Z_l = \frac{X-\mu}{\sigma} = \frac{37.9-38.5}{0.65} = -0.92\,; Z_u = \frac{39.8-38.5}{0.65} = 2$

Look up probabilities

$p(z \geq Z_l) = p(z \geq -0.92) = p(z \leq 0.92)$

$p(z \geq Z_l) = 1 - p(z \geq 0.92) = 1 - 0.1788 = 0.8212\,(exact\ value\ 0.7993)$

$p(z \geq Z_u) = p(z \geq 2.0) = 0.0228$

$p(37.9 \leq T \leq 40.1)$=0.8212-0.0228 = 0.7984 (exact value 0.7993)

(ii) A border control agency decides to randomly screen 5 passengers for disease X from each incoming flight.

a. Using the notation that a particular flight has N passengers of whom I are infected, write down an expression for the probability that the group selected from one flight contains exactly i infected passengers. Include any conditions on your parameters.

b. Of two incoming flights, flight F1 has 50 passengers of whom 3 are infected and flight F2 has 80 passengers of whom 6 are infected. Is it more likely that at least one infected passenger will be selected from flight F1 or flight F2? (probabilities must be shown for credit).

Answer:

(a) Order not important $p(¿infected = i) = \frac{I\,C_i \times (N-I)\,C_{(5-i)}}{N\,C_5}$ where $0 \leq i \leq I$

(b) $p(at\ least\ 1\ on\ F1, N=50, I=3) = 1 - p¿$

$¿1 - p(i=0) = 1 - \frac{3\,C_0 \times 47\,C_{5-0}}{50\,C_5} = 1 - 0.724 = 0.276$

$p(at\ least\ 1\ on\ F2, N=80, I=6) = 1 - p¿$

$¿1 - p(i=0) = 1 - \frac{6\,C_0 \times 74\,C_{5-0}}{80\,C_5} = 1 - 0.67 = 0.33$

Therefore, more likely at least one infected passenger will be selected from flight 2.

(i) Any passenger with a temperature 38°C or higher is recorded as potentially infected with disease X. 78% of infected individuals with disease X will have a temperature of at least 38°C. However, 5% of all arriving passengers who do not have the disease will also have a temperature of at least 38°C. Border control keeps a list of the prevalence of the disease in each country. A passenger arriving from Country Z, which has a reported prevalence of disease X of 1.5%, has a temperature of 38.7°C.

a. What is the probability, given the information available to border control, that the passenger is infected with disease X?

Answer:

$$p\left(disease\ X \middle| +ve\right) = \frac{p\left(HighTemp \middle| X\right) p\left(X\right)}{p\left(HighTemp \middle| X\right) p\left(X\right) + p\left(HighTemp \middle| \;\; X\right) p\left(\;\; X\right)} = \frac{0.78 \times 0.015}{0.78 \times 0.015 + 0.05 \times 0.985} = 0.192$$

b. The passenger is then asked to provide a swab for a lateral flow test. Suppose the lateral flow test has a specificity of 99% and sensitivity of 70%. The passenger tests positive. What is the probability, given all the information available to border control, that the passenger in infected with disease X?

(recall: specificity – the proportion of people without the disease that have a negative test; sensitivity – the proportion of people with the disease that have a positive test).

Answer: updated prior probability disease X = 0.192 (from part a)

$$p\left(diseaseX \middle| +veLFT\right) = \frac{p\left(+veLFT \middle| X\right) p\left(X\right)}{p\left(+veLFT \middle| X\right) p\left(X\right) + p\left(+veLFT \middle| \;\; X\right) p\left(\;\; X\right)} = 0.943$$

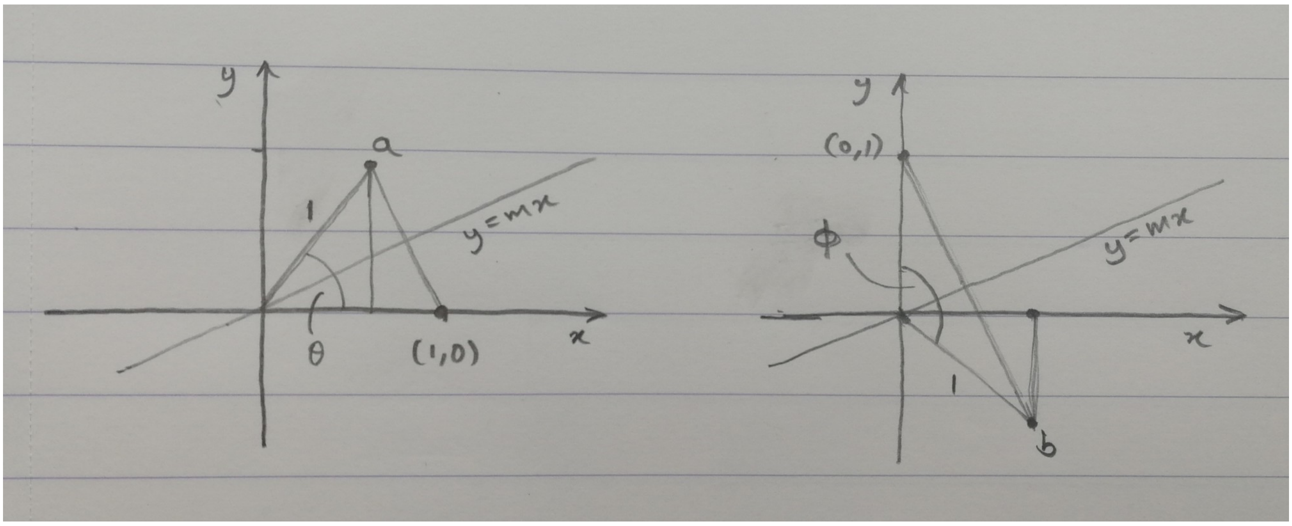
A2

(i) Consider the linear geometric transformation of points in the two-dimensional plane whereby points are reflected in the line $y = mx$, where m is a scalar constant.

Construct a 2 x 2 matrix A representing this transformation.

Answer:

The columns of the transformation matrix are the transformed positions of the basis vectors. Taking basis (0, 1) and (1, 0), the following sketch shows the relevant geometry:



Point a is the transformed location of (1, 0), and point b is the transformed position of (0, 1). Angle θ is given by $\tan\theta = m$, while $\phi = \pi/2 - \theta$.

Thus a = $(\cos 2\theta, \sin 2\theta)$, while b = $(\cos(\pi/2 - 2\theta), -\sin(\pi/2 - 2\theta)) = (\sin 2\theta, -\cos 2\theta)$.

Hence a form of the matrix is:

$$M = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

Using the double angle formulae:

$\cos 2\theta = (1 - \tan^2\theta) / (1 + \tan^2\theta)$

$\sin 2\theta = 2\tan\theta / (1 + \tan^2\theta)$

we can write this as

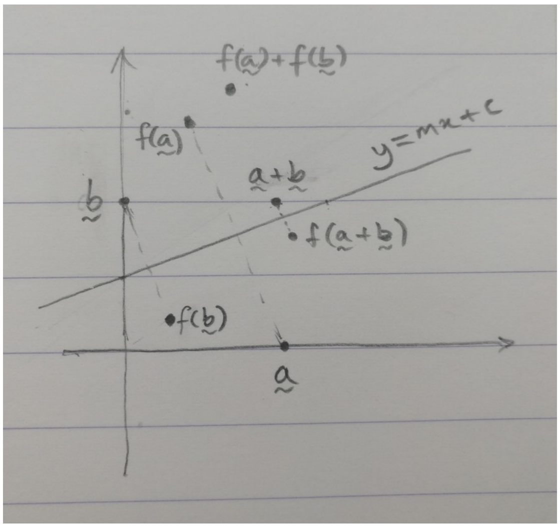
$$M = \frac{1}{1+m^2}\begin{pmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{pmatrix}$$

(extra credit for showing this).

(ii) Show, by sketching a suitable example or otherwise, that reflection in a line $y = mx + c$, where c is nonzero, is **not** a linear transformation of points in the two-dimensional plane.

Answer:

Any suitable sketch showing that for two vectors ***a*** and ***b***, the sum of their transformed positions f(***a***) + f(***b***) is not the same as their transformed sum f(***a*** + ***b***), e.g.



(iii) Find a complete set of eigenvalues and eigenvectors for the following matrix:

$$M=\begin{pmatrix}3 & 1 & 1\\ 2 & 4 & 2\\ 1 & 1 & 3\end{pmatrix}$$

Answer:

The characteristic equation reduces to:

$$det(M-\lambda I)=\lambda^3-10\lambda^2+28\lambda-24=(6-\lambda)(2-\lambda)^2$$

This has roots at 6 and 2 (a double root).

For $\lambda_1 = 6$ we have $(M - 6I)\mathbf{v}_1 = \mathbf{0}$, which reduces to

$x + y - 3z = 0$

$-y + 2z = 0$

and hence

$$\boldsymbol{v}_1=\alpha\begin{pmatrix}1\\ 2\\ 1\end{pmatrix}$$

(or equivalent) for any scalar α.

For $\lambda_2 = 2$ we have $(M - 2I)\mathbf{v}_2 = \mathbf{0}$, which reduces to

$x + y + z = 0$

and hence

$$\boldsymbol{v}_1=\alpha\begin{pmatrix}1\\ -1\\ 0\end{pmatrix}+\beta\begin{pmatrix}1\\ 0\\ -1\end{pmatrix}$$

(or equivalent) for any scalars α and β.

B3

(a) A population ecologist wants to investigate the factors that favour the extinction of lizard populations. She uses a long term monitoring dataset, and quantifies how many populations have gone extinct and how many are still extant. She also divides populations as being viviparous (giving birth to live young) or oviparous (laying eggs). When tabulating the results, she gets:

	Extinct	Surviving
Viviparous	16	74
Oviparous	5	81

Test whether reproductive mode affects the probability of extinction of a population? *[~20% marks]*

This is an intrinsic chisquare test. We first compute the individual expected values as (row_sum * col_sum)/total

	Extinct	Surviving
Viviparous	10.7	79.3
Oviparous	10.2	75.7

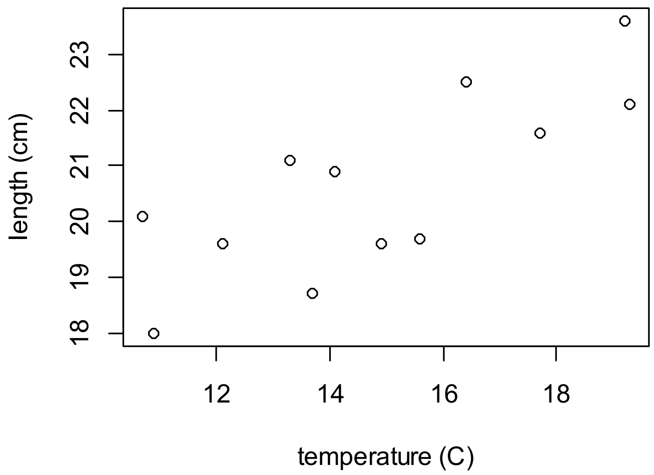
The chisquare value is the sum of $(E-O)^2/E$, which is 5.99.
We compare our computed chisq value of 5.99 with 1df against the critical value in the table. As the observed value is larger than the critical value, we conclude that we can reject the H0 (there is a significant link between reproductive mode and extinction probability)

(b) Provide a sentence that you could use to report the results in a scientific paper.

[~10% marks]

Viviparous populations were more likely to go extinct than oviparous ones (Chisq=5.99, df=1, $p<0.05$)

(c) Our population ecologist is also interested in whether local temperature is linked to size of lizards at that location. She collects lizards from 12 locations, and then plots length (cm) against the mean summary temperature (degrees Celsius):



She runs a linear regression to test whether length is predicted by temperature. Complete the following ANOVA table, and test whether the relationship is significant:

	SS	df	MS	F
Model	17.86	1	17.86	15.2
Residual	11.76	10 (n-2)	1.176	
Total	29.62			

From the table, the critical value for F1,10 is smaller than the observed F, we have evidence to reject the null hypothesis, so there is a significant relationship between length and temperature

[~30% marks]

(d) Provide a sentence to report your results in a paper

[~10% marks]

Lizard length increased with increasing mean summer temperature ($F_{1,10}$=15.2, p<0.05)

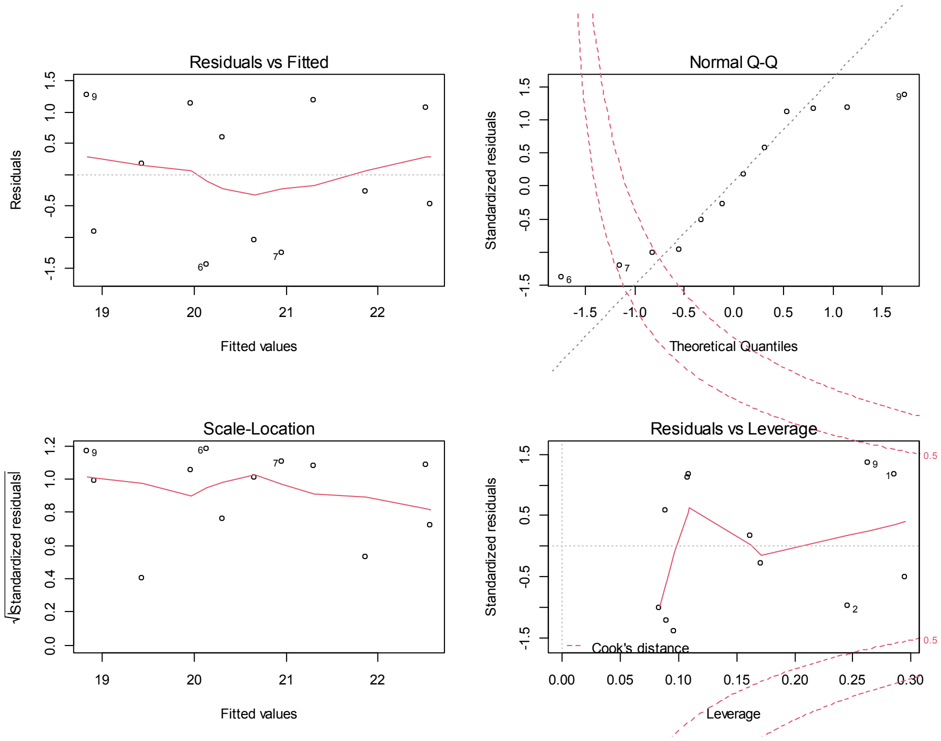
(e) Given that the above analysis gives a line of best fit with an intercept of 14.18 and a slope of 0.43, what would be the predicted length of a lizard collected from a location with a mean temperature of 15.2 degrees? And one for a location with a temperature of 21.5 degrees? Which prediction is likely to be more accurate?

[~10% marks]

Length for 15.2 C: 14.18+0.43*15.2=20.7 cm
Length for 21.5 C: 14.18+0.43*21.5=23.4 cm

From the plot above, we can see that, at 21.5 C, we are extrapolating beyond the observed range of temperature. Thus, this estimate is less likely to be accurate than the one for 15.2 C, which is in the middle of the observed range.

(f) Comment on the following diagnostic plots

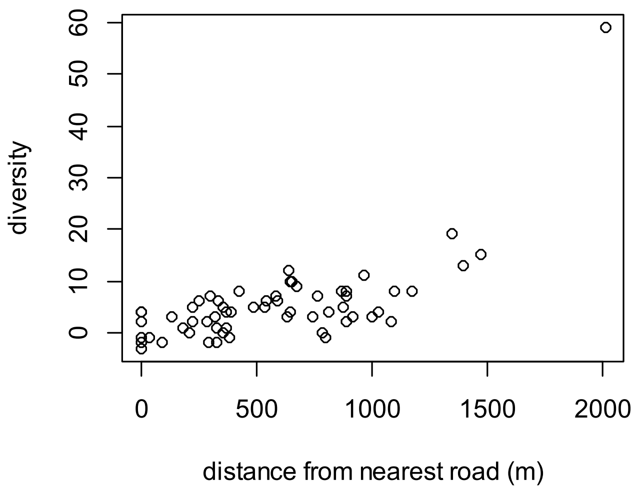


[~20% marks]

TopLeft graph: We meet the expectation of an even cloud of points above and below zero.
BottomLeft graph: We meet the expectation of an event cloud of points without extreme values on the y axis
TopRight: We expected a linear relationship on the QQ plot, but there is a clear deviation. We might need to transform our data, as the residuals are NOT normal.
BottomRight: expect all points within the Cook's isocline, and they are well within those limits.

B4

(a) A conservation biologist wants to test whether distance from a road affects bird biodiversity found at given spot. He runs 60 surveys, and plots bird diversity (number of species observed during a survey) versus distance from the nearest road (m).



He tests whether distance predicts diversity, and gets:

```
Call:
lm(formula = count ~ distance, data = birds)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.430155   1.351134  -1.799   0.0773 .
distance     0.013409   0.001893   7.085 2.11e-09 ***
```

Given that the $SS_{Residual}$ is 2192 and SS_{Total} is 4089, test whether there is a significant effect of distance

[~30% marks]

We first build an ANOVA table:

	SS	df	MS	F
Model	1897	1	1897	50.2
Residual	2192	58 (n-2)	37.79	
Total	4089			

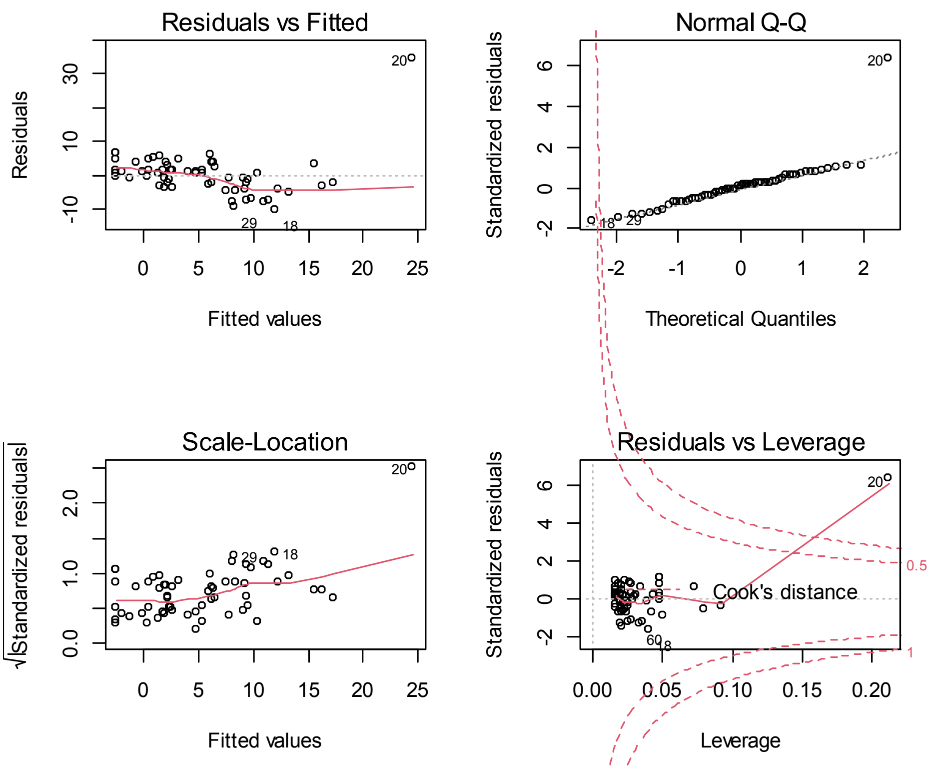
As the computed F for 1,58 degrees of freedom is much larger than the critical values from the table, we conclude that the relationship is significant.

(b) Provide a sentence that you could use to report the results in a scientific paper.

Bird diversity increased significantly with increasing distance from the nearest road ($F_{1,58}$=50.2, p<0.05).

[~10% marks]

(c) Here are the diagnostic plots:



What can you conclude?

All four plots reveal problems that indicate the assumptions were violated:
TL plot: expect even cloud above and below zero, we note that residuals initially dip below zero consistently, and then point 20 is a very large positive value (i.e. assumption of homogeneity of variances not met)
BL plot: expect an even cloud without points sticking out, but point 20 is far out on the y axis (i.e. assumption of homogeneity of variances not met)
TR plot: expect 1:1 relationship, point 20 deviates strongly (we fail to meet the normality of residuals assumption)
TB plot: expect points within 0.5 isocline, but point 20 is outside the 1 isocline of Cook's distances (i.e. it has a large impact on the estimate relationship

[~10% marks]

(d) He then attempts to fit a General Linear Model with a Poisson Error structure, and gets the output

```
Call:
glm(formula = count ~ distance, family = "poisson", data = birds)

Deviance Residuals:
```

```
    Min        1Q    Median        3Q       Max
-2.5233   -0.8336   0.0171   0.4465   2.1981

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.787e+00  8.992e-02  19.870   <2e-16 ***
distance    -1.726e-05  1.393e-04  -0.124    0.901
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 64.936  on ??  degrees of freedom
Residual deviance: 64.921  on ??  degrees of freedom
```

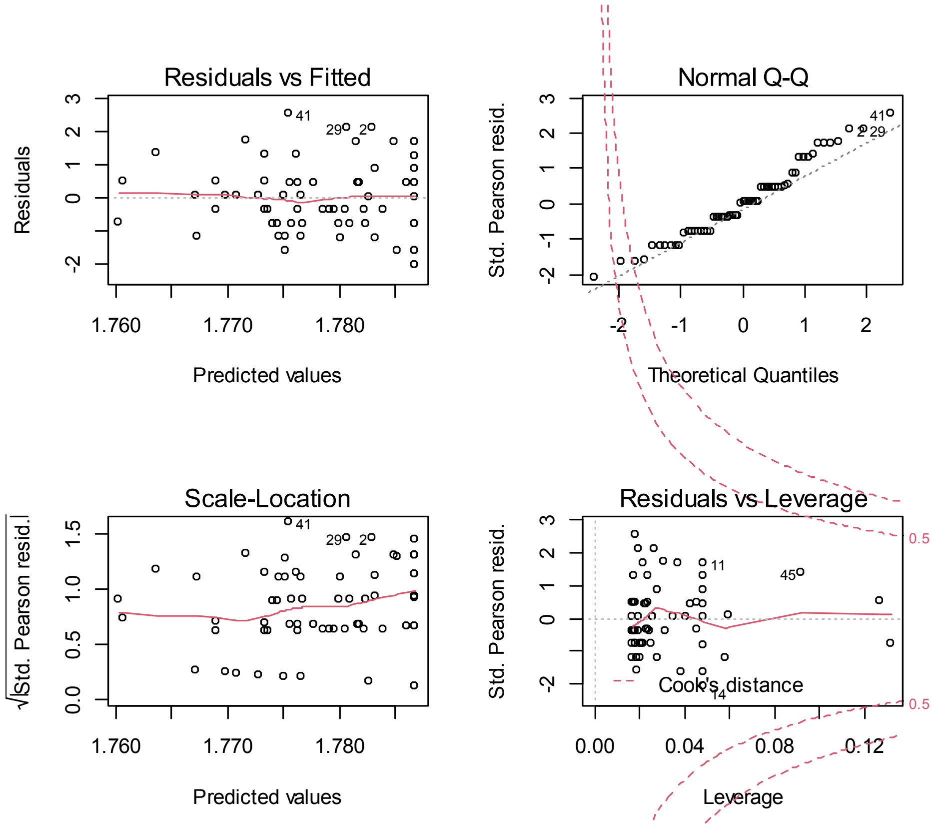
Compute the df, and then test whether the relationship is significant.

[~20% marks]

Null deviance = N-1 = 60-1 = 59
K = number of coefficients in model = 2 (slope and intercept)
Res deviance = N-K= 58

Estimate chisquare: delta deviance given delta df: 64.936-64.921=0.015 with 59-58=1 degrees of freedom.
As this value is smaller than the critical value in the table, there is not sufficient evidence to reject the H_0

(e) Here are the diagnostic plots



The issues present in (c) have been resolved for each plot.

How do these plots compare to the one for the linear model in (c)?

[~10% marks]

(f) Use the relationships estimated in (a) and (d) to predict bird diversity at 1900 meters from the nearest road. Can you explain the discrepancy between these predictions

[~20% marks]

-2.430155+0.013409*1900=23.0

exp(1.787e+00-1.726e-05*1900)=5.8 (remembering to use the exp to rescale the prediction)

The first analysis was strongly driven by an outlier, which had very high leverage, and so this leads to inflated predictions from the linear regression

MB 2021 C5 Solution

a) Rate at which infected individuals lose infectivity (note: no credit for "death" alone).

b) (Lightly adapted) Bookwork: $I=I_0\exp ¿¿$

c) $$\frac{d\rho}{dt}=\frac{(I_1+I_2)\frac{dI_2}{dt}-I_2\left(\frac{dI_1}{dt}+\frac{dI_2}{dt}\right)}{(I_1+I_2)^2}\,\frac{d\rho}{dt}=\frac{I_1\frac{dI_2}{dt}-I_2\frac{dI_1}{dt}}{(I_1+I_2)^2}\,\frac{d\rho}{dt}=\frac{I_1I_2\big((\alpha_2N-\mu)-(\alpha_1N-\mu)\big)}{(I_1+I_2)^2}$$

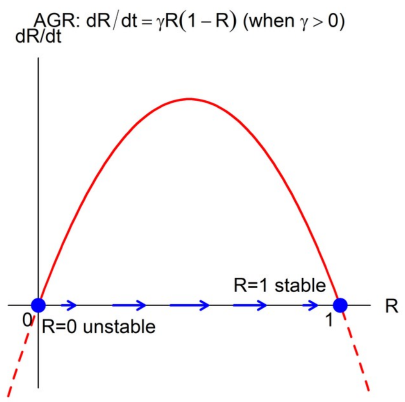
$$\frac{d\rho}{dt}=(\alpha ¿¿ 2-\alpha_1)N\frac{I_2}{I_1+I_2}\frac{I_1}{I_1+I_2}¿$$

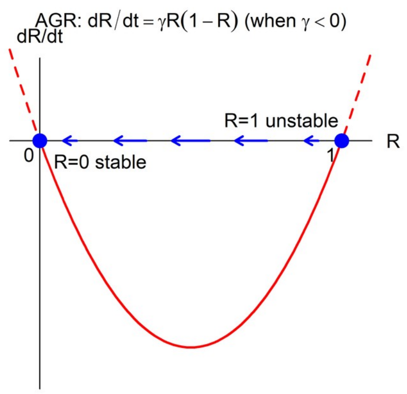
Since $\rho=\frac{I_2}{I_1+I_2}$ and so $\frac{I_1}{I_1+I_2}=1-\rho$, the result follows using the definition of γ.

d) Equilibria when $\frac{d\rho}{dt}=F(\rho)=\gamma\rho(1-\rho)=0$, i.e. $\rho=0$ or $\rho=1$.

Stability via $\frac{dF}{d\rho}=\gamma(1-2\rho)$ evaluated at each equilibrium, i.e. $\frac{dF}{d\rho}(0)=\gamma$ and $\frac{dF}{d\rho}(1)=-\gamma$. It therefore depends on sign of $\gamma=(\alpha_2-\alpha_1)N$.

- Case i. $\alpha_2>\alpha_1$. Strain 2 has a selective advantage because it infects more quickly, and so would be expected to eventually spread to fixation
  - Here $\gamma>0$, and so $\rho=0$ is unstable (since $\frac{dF}{d\rho}(0)>0¿$, $\rho=1$ is stable (since $\frac{dF}{d\rho}(1)<0¿$
- Case ii. $\alpha_1>\alpha_2$. Strain 2 has a selective disadvantage, and so would be expected to be eliminated in the long term
  - Here $\gamma<0$ and correspondingly $\rho=0$ is stable, $\rho=1$ unstable.
- Case iii. $\alpha_2=\alpha_1$. The strains are epidemiologically neutral
  - Here $\gamma=0$ and so every value of ρ is an equilibrium; the proportion of cases due to variant 2 remains fixed at $\rho=\rho_0$ forever.





e) (Lightly adapted) Bookwork

$$\rho = \frac{1}{1 + \left(\frac{1-\rho_0}{\rho_0}\right)\exp\left(-\gamma t\right)}$$

In all cases sketches should show a very low initial value of ρ, and should follow the cases outlined in part (d), i.e. $\alpha_2 > \alpha_1$ shows sigmoidal growth up to $\rho = 1$; $\alpha_1 > \alpha_2$ shows decrease down to $\rho = 1; \alpha_1 = \alpha_2$ shows no change, with $\rho = \rho_0$ forever.

f) Using the result of part (b) leads to $I_i = I_{0i}\exp\text{¿¿}$. Substituting these directly into the definition of ρ leads to

$\rho = I_{02}\exp\text{¿¿¿}$ $\rho = \frac{1}{\left(\frac{I_{01}}{I_{02}}\right)\exp\text{¿¿¿¿}}$ $\rho = \frac{1}{\left(\frac{I_{01}}{I_{02}}\right)\exp\left(\left(\alpha_1 - \alpha_2\right)Nt\right) + 1}$ But noting the definition of $\gamma = \left(\alpha_2 - \alpha_1\right)N$ and that

$$\frac{1-\rho_0}{\rho_0} = \frac{1 - \frac{I_{02}}{I_{01}+I_{02}}}{\frac{I_{02}}{I_{01}+I_{02}}} = \frac{\frac{I_{01}}{I_{01}+I_{02}}}{\frac{I_{02}}{I_{01}+I_{02}}} = \frac{I_{01}}{I_{02}}$$

shows this is precisely what we obtained in part (e).

g) Give any sign of life some credit here. The model predicts the proportion of any new variant increases logistically if it has a selective advantage. So any mention of comparing survey data with the results of the model in an effort to estimate γ gets some marks.

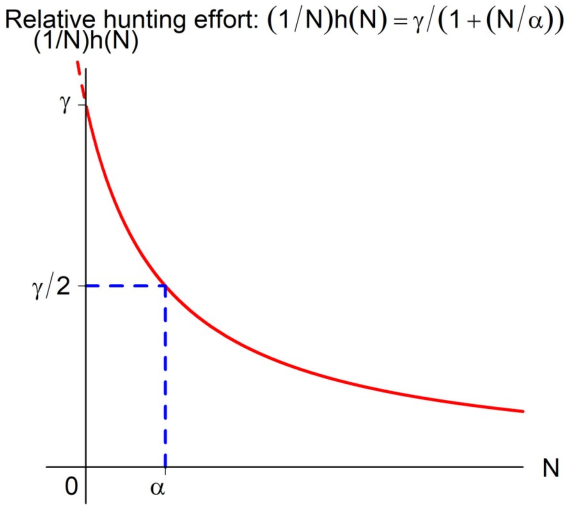
But – at least until the scripts come in, since this might end up being altered in the light of what turns up – for full credit need either:

- EITHER: a somewhat sketched out way of estimating γ (e.g. nonlinear least squares, or linear regression of $logit(\rho)$ upon t) even if highly error prone (e.g. any calculation based on only two values of ρ)
- OR: a clear description of the idea that, to quantify the relative infectivity (rather than the difference in growth rates), need to have some notion of the underlying growth rate

(Note for supervisors: the basic set up of this question is lifted from the models reported in the paper https://www.medrxiv.org/content/10.1101/2020.12.24.20248822v2.full.pdf)

MB 2021 C6 Solution

a) $h(N)/N$ is the relative hunting effort, i.e. how the rate of hunting <u>per individual</u> depends on the population size.



Since $h(N)/N$ goes up as N gets smaller, this is consistent with the idea of the hunters targeting a species more when it is rarer and so more prized by collectors.

b) The condition $\gamma > \beta$ means that at very low population sizes the relative rate of death due to hunting exceeds the net relative population growth rate. This is clearly a pre-condition for the species to be driven extinct by hunting.

c) At equilibrium

$$0 = \beta N(1-N) - \frac{\gamma N}{1+(N/\alpha)}$$

and so

$$0 = \beta N\left((1-N) - \frac{\alpha\gamma/\beta}{\alpha+N}\right)$$

There is therefore always an equilibrium when $N=0$.

Any other equilibria must satisfy

$$(1-N)(\alpha+N) - \frac{\alpha\gamma}{\beta} = 0 \quad -N^2-(\alpha-1)N-\left(\frac{\alpha\gamma}{\beta}-\alpha\right)=0 \quad N^2-(1-\alpha)N+\alpha\left(\frac{\gamma}{\beta}-1\right)=0$$

d) *(The form of the quadratic to aim for makes the analogy with $\lambda^2 - T\lambda + \Delta = 0$ very explicit, so some might draw a parallel with classifying stability of a linear system and – correctly – reason that two positive roots corresponds to the case of an unstable focus, i.e. $T=(1-\alpha)>0$, $\Delta = \alpha\left(\frac{\gamma}{\beta}-1\right)>0$ and $T^2 - 4\Delta = (1-\alpha)^2 - 4\alpha\left(\frac{\gamma}{\beta}-1\right)>0$. That is the expected method).*

However, without seeing this, can argue that real solutions need a positive discriminant, i.e.

$$(1-\alpha)^2 > 4\alpha\left(\frac{\gamma}{\beta}-1\right)$$

Since the solutions are

$$N=\frac{1}{2}\left((1-\alpha)\pm\sqrt{(1-\alpha)^2-4\alpha\left(\frac{\gamma}{\beta}-1\right)}\right)$$

For both to be positive we clearly require $\alpha<1$ and $\frac{\gamma}{\beta}>1$

e) Differentiating $h(N)$ leads to

$$\frac{dh}{dN}=\frac{\left(1+\left(\frac{N}{\alpha}\right)\right)\gamma-\left(\frac{1}{\alpha}\right)\gamma N}{\left(1+\left(\frac{N}{\alpha}\right)\right)^2}=\frac{\gamma}{\left(1+\left(\frac{N}{\alpha}\right)\right)^2}$$

Since $g(N)=f(N)-h(N)$

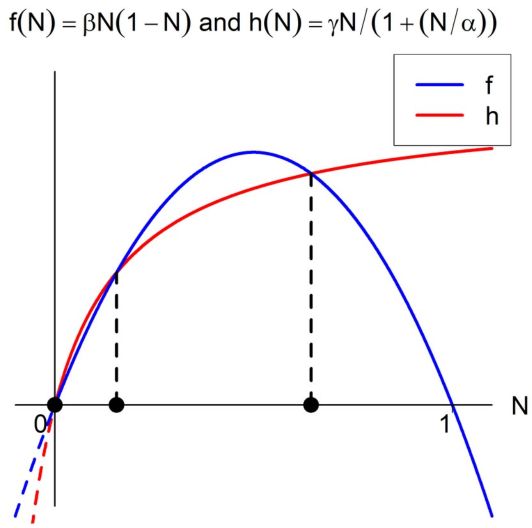
$$\frac{dg}{dN}=\beta(1-2N)-\frac{\gamma}{\left(1+\left(\frac{N}{\alpha}\right)\right)^2}$$ So

$$\frac{dg}{dN}(0)=\beta-\gamma$$

which is negative since $\gamma > \beta$. The equilibrium at $N=0$ is therefore stable.

The graphs which follow were calculated in the particular case $\alpha=\frac{1}{5}, \beta=\frac{1}{2}$ *and* $\gamma=\frac{3}{4}$

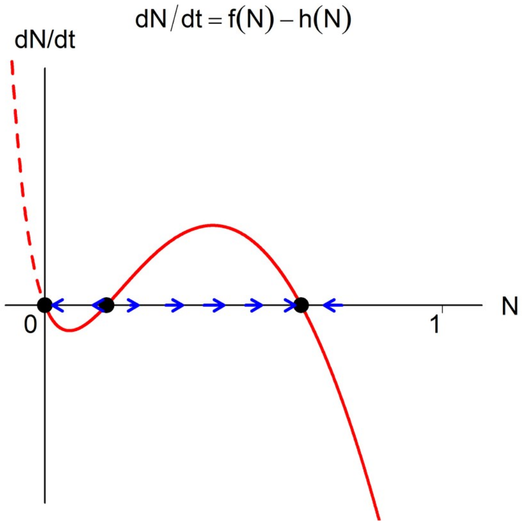
f)



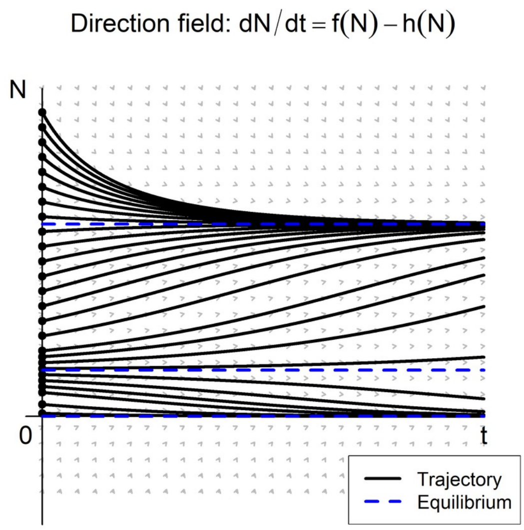
g) The lower positive equilibrium is unstable; the upper one stable. Students need to write some text to justify this, e.g. For the largest equilibrium

- if N is increased from the equilibrium, then $\frac{dN}{dt}<0$ and so N decreases
- if N is decreased from the equilibrium, then $\frac{dN}{dt}>0$ and so N increases

and so the dynamics of the system oppose any perturbation, meaning the equilibrium is stable.



h) If the initial condition is too low, the population dies out; there is an Allee effect. If the initial condition is high, then the population does tend to a positive equilibrium, but with $N<1$ (i.e. at a smaller density than it would be without hunting).



The Allee effect is driven by the hunting effort increasing the rarer the focal species becomes. It can therefore be driven to extinction if its density ever goes below a threshold

(Note for supervisors: the basic set up of this question is lifted from the models reported in the paper https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0040415)

Mathematical Biology Exam 2021

D7

a) $\begin{cases} \dfrac{dx}{dt} = xy+3 = 0 \\ \dfrac{dy}{dt} = xy - x^2 - 4x = 0 \end{cases} \Leftrightarrow \begin{cases} y = -3/x \quad (1) \\ x(y-x-4) = 0 \\ x=0 \text{ OR } y = x+4 \\ (2a) \qquad (2b) \end{cases}$

(1) + (2a) : $xy = -3$

$0 = -3$ ✗ no solutions

(1) + (2b) : $y = -3/x$

$x+4 = -3/x$

$x^2+4x+3 = 0$

$(x+1)(x+3) = 0$

$x=-1$ OR $x=-3$

$\Downarrow \qquad \Downarrow$

$y = -3/-1 = 3 \qquad y = -3/-3 = 1$

$(-1,3)$ and $(-3,1)$ are equilibrium points

$x' = xy+3 \quad , \quad y' = xy - x^2 - 4x$

	$(-1,3)$	$(-3,1)$

D_{11}	y	3	1
D_{12}	x	-1	-3
D_{21}	$y-2x-4$	1	3
D_{22}	x	-1	-3
T		$3-1=2$	$1-3=-2$
Δ		$-3+1=-2$	$-3+9=6$
$T^2-4\Delta$		—	$(-2)^2-4(6)=4-24=-20<0$
		Saddle point	Stable focus

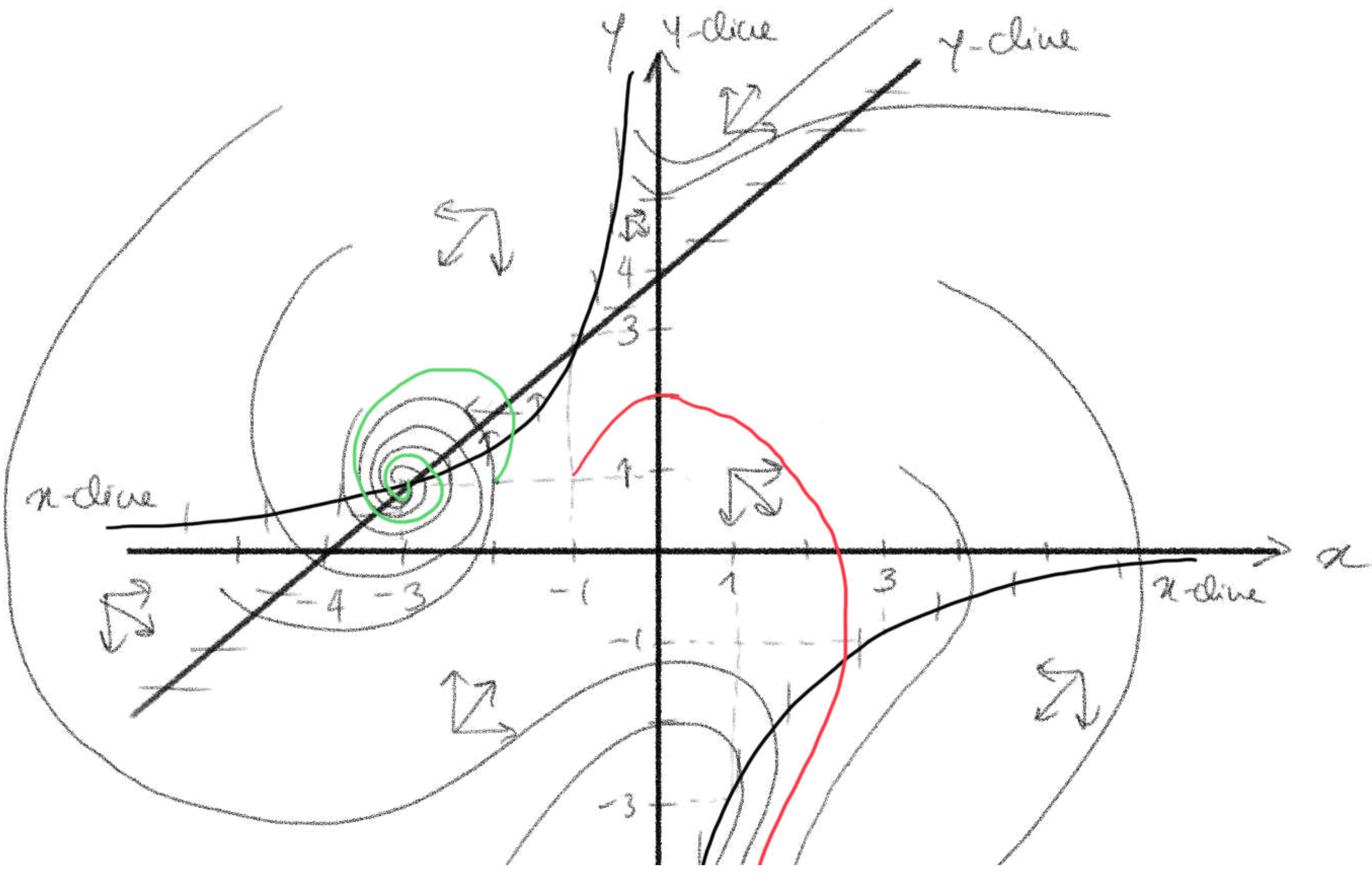
b)

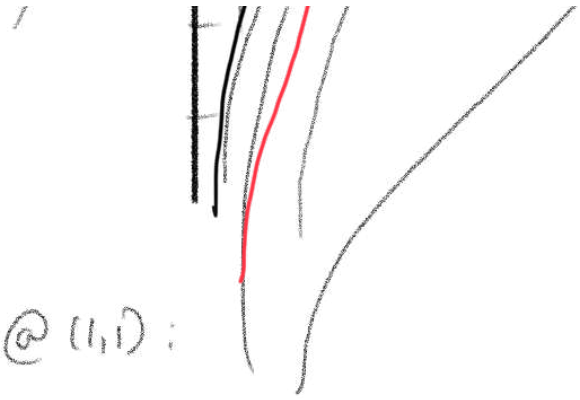
x-nullcline

$y=-3/x$

y-nullclines

$x=0$ or $y=x+4$



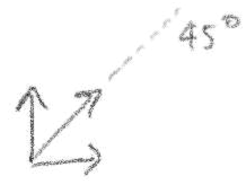


Direction field @ (1,1):

$x' = 1.1+3 = 4$

$y' = 1.1 - 1^2 - 4(1) = -4$

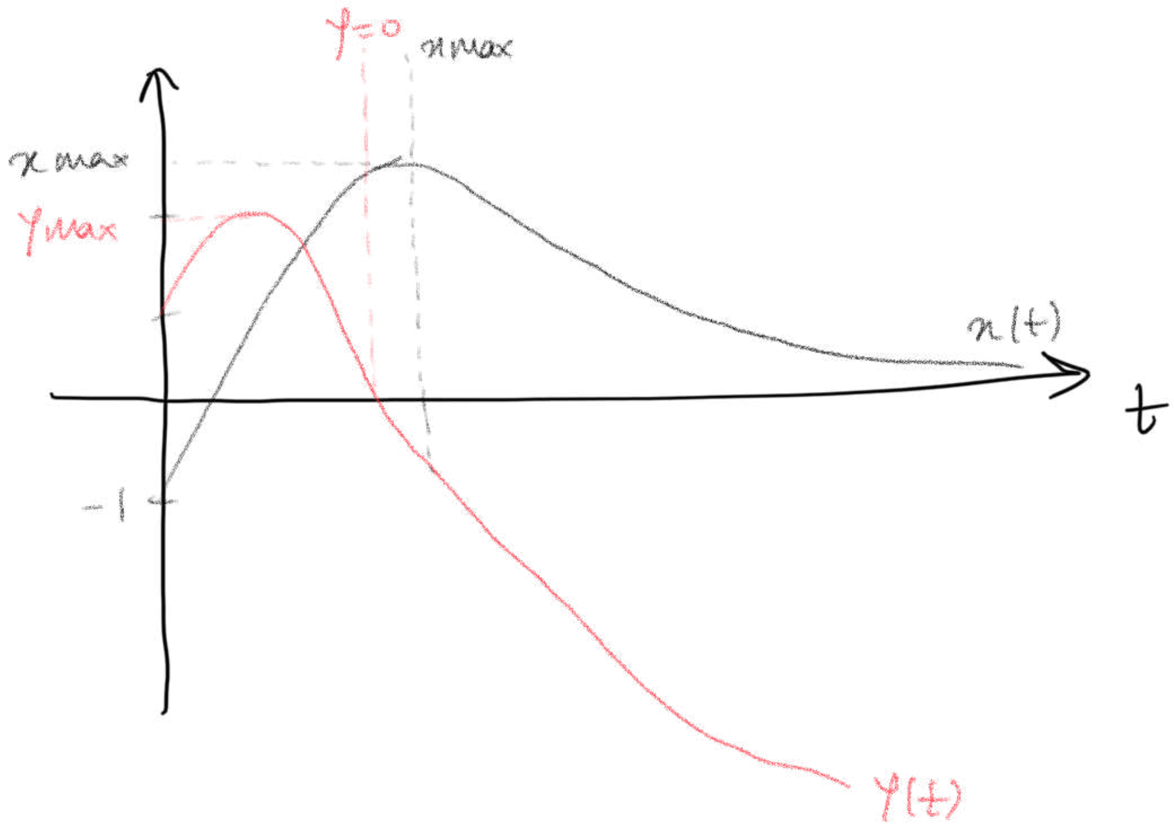
c) $x(0) = -1$, $y(0) = 1$



@ $t=0$: $x'(0) = (-1).1+3 = -1+3 = 2$

$y'(0) = (-1).1 - (-1)^2 - 4(-1) = -1-1+4 = 2$

— trajectory for (-1,1)



d) Direction field @ (-2,1) :

$x' = -2(1)+3 = -2+3 = 1$

$y' = -2(1) - (-2)^2 - 4(-2) = -2-4+8 = -6+8 = 2$

— trajectory for (-2,1)

For $(x_0, y_0) = (-1, 1)$ the direction field @ $t=0$ is $x' = y' = 2$ which means initially the trajectory will be moving equally fast along both directions and it will approach the nullcline $x=0$, while moving away from the other nullclines. For $(x_0, y_0) = (-2, 1)$, however, the direction field @ $t=0$ is $x' = 1$ and $y' = 2$, which means the trajectory initially moves twice as fast along y and therefore it will reach the x-nullcline $y = -3/x$ first; once it crosses this nullcline the trajectory will be captured by the stable focus @ $(x, y) = (-3, 1)$ leading to long term solutions $x \to -3$ and $y \to 1$, which are noticeably different from the long term solutions in (c) $(x \to 0,\ y \to -\infty)$.

Mathematical Biology 2020-21
Exam question from Olivier Restif (Block D)

Two strains of *E. coli* bacteria, labelled *X* and *Y*, are grown together in liquid broth. The differential equations below represent the dynamics of the densities of *X* and *Y* respectively:

$$\frac{dX}{dt} = \beta\, X\, \left(1 - \frac{X + \alpha\, Y}{K}\right)$$
$$\frac{dY}{dt} = \beta\, Y\, \left(1 - \frac{X + Y}{K}\right)$$

a. Give an interpretation for each of the three parameters β, α and K, assuming they can only take positive values. [15%]
b. Calculate and sketch the nullclines, stationary points and direction field in the (X,Y) phase plane, considering two alternative cases:
   - Case 1: α < 1;
   - Case 2: α > 1. [50%]
c. Determine the stability of the stationary points in Case 1. [35%]

Model answers (each red dot represents 1 mark, for a total of 20).

a.
- β is the intrinsic growth rate per capita of either strain on its own.
- K is the carrying capacity of either strain on its own.
- α is the relative effect of strain Y on the growth of strain X (inter-strain competition).

b.
(i) X nullclines:
- dX/dt = 0 has two solutions: X=0 or Y = (K-X)/α

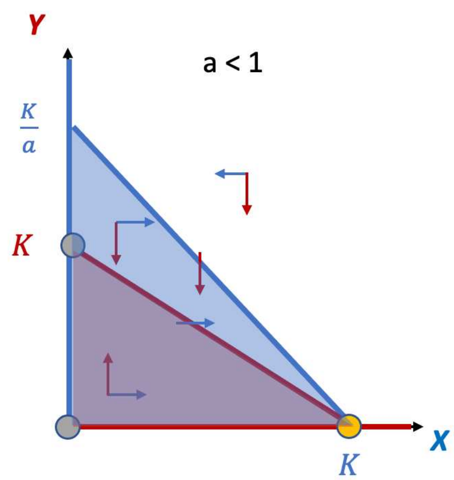
(ii) Y nullclines:
- dY/dt = 0 has two solutions: Y=0 or Y=K-X

(iii) In Case 1 or case 2, there are three stationary points, assuming α ≠ 1:
- (X=0,Y=0)
- (X=0, Y=K)
- (X=K, Y=0)

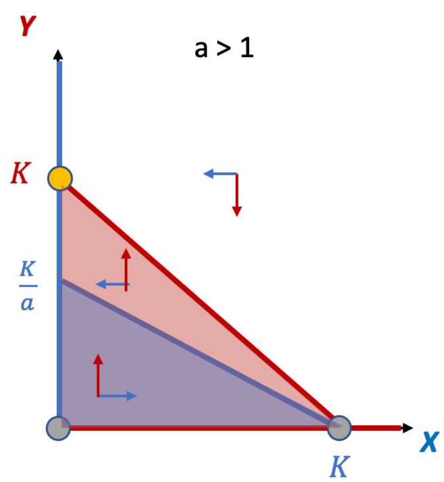
(iv) Direction field:
- near the origin, dX/dt ≈ β X > 0 and dY/dt ≈ β Y > 0.

(v) Phase plane:

••



••



c. Here we assume α < 1.
i. Let

$$F(X,Y) = \beta\, X\left(1 - \frac{X + \alpha\, Y}{K}\right) = \beta X - \frac{\beta X^2}{K} - \frac{\beta \alpha X Y}{K}$$
$$G(X,Y) = \beta\, Y\left(1 - \frac{X + Y}{K}\right) = \beta Y - \frac{\beta Y^2}{K} - \frac{\beta X Y}{K}$$

- $D_{1,1} = \frac{\partial F}{\partial X} = \beta - \frac{2\beta X}{K} - \frac{\beta\alpha Y}{K}$,
- $D_{1,2} = \frac{\partial F}{\partial Y} = -\frac{\beta\alpha X}{K}$
- $D_{2,1} = \frac{\partial G}{\partial X} = -\frac{\beta Y}{K}$,
- $D_{2,2} = \frac{\partial G}{\partial Y} = \beta - \frac{2\beta Y}{K} - \frac{\beta X}{K}$

ii. (X=0,Y=0):

$$D_{1,1} = \beta\,, \qquad D_{1,2} = 0$$
$$D_{2,1} = 0\,, \qquad D_{2,2} = \beta$$
$$T = 2\beta > 0$$
$$\Delta = \beta^2 > 0$$

- => Unstable

iii. (X=K, Y=0)

$$D_{1,1} = -\beta\,, \qquad D_{1,2} = -\beta\alpha$$
$$D_{2,1} = 0\,, \qquad D_{2,2} = 0$$
$$T = -\beta < 0$$
$$\Delta = 0$$

- => Stable

iv. (X=0, Y=K)

$$D_{1,1} = \beta(1-\alpha)\,, \qquad D_{1,2} = 0$$
$$D_{2,1} = -\beta\,, \qquad D_{2,2} = -\beta$$
$$T = -\beta\alpha < 0$$
$$\Delta = -\beta^2(1-\alpha) < 0$$

- => Saddle point.

MB 2021 E9 (SOLUTIONS)

The following scheme describes the mechanism of a fully reversible enzyme catalysed reaction, in which the enzyme, E, facilitates the conversion of a single substrate, S, into a product, P, via the formation of an enzyme-substrate complex, ES.

$$\mathrm{E + S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{E + P}$$

(a) Construct a mass balance equation for the total concentration of enzyme that is present at time zero, $[E]_0$.

[~5% marks]

$[E]_0 = [E] + [ES]$

(b) Write down two differential equations that describe how the rates of change of [ES] and [P] vary as a function of time, *t*.

[~10% marks]

$d[ES]/dt = k_1[E][S] - k_2[ES] - k_3[ES] + k_4[E][P]$

$d[P]/dt = k_3[ES] - k_4[E][P]$

(c) Using your results from part ***(b)***, apply the steady state approximation to the enzyme-substrate complex and construct an expression for [ES] in terms of rate constants, the current concentration of free enzyme, [E], the current concentration of substrate, [S], and the current concentration of product, [P].

[~10% marks]

$d[ES]/dt = 0 = k_1[E][S] - k_2[ES] - k_3[ES] + k_4[E][P]$

i.e. $[ES](k_2 + k_3) = k_1[E][S] + k_4[E][P]$

i.e. $[ES] = k_1[E][S]/(k_2 + k_3) + k_4[E][P]/(k_2 + k_3)$

(d) Using your results from part ***(c)***, construct an expression for [ES] in terms of rate constants, the total concentration of enzyme present at time zero, $[E]_0$, the current concentration of substrate, [S], and the current concentration of product, [P].

[~10% marks]

$[ES] = k_1\{[E]_0 - [ES]\}[S]/(k_2 + k_3) + k_4\{[E]_0 - [ES]\}[P]/(k_2 + k_3)$

i.e. $[ES](k_2 + k_3) = k_1[E]_0[S] - k_1[ES][S] + k_4[E]_0[P] - k_4[ES][P]$

i.e. $[ES](k_1[S] + k_2 + k_3 + k_4[P]) = [E]_0\{k_1[S] + k_4[P]\}$

i.e. $[ES] = [E]_0\{k_1[S] + k_4[P]\}/(k_1[S] + k_2 + k_3 + k_4[P])$

(e) Construct an expression for the rate of change of [P] in terms of rate constants, the total concentration of enzyme present at time zero, $[E]_0$, the current concentration of the enzyme-substrate complex, [ES], and the current concentration of product, [P].

[~5% marks]

$$d[P]/dt = k_3[ES] - k_4[E][P] = k_3[ES] - k_4\{[E]_0 - [ES]\}[P]$$

(f) Using your results from parts ***(d)*** and ***(e)***, show that V, the rate of formation of product, obeys the following rate law equation:

$$V = \frac{\dfrac{V_{MAX}{}^{f}}{K_M{}^{f}}[S] - \dfrac{V_{MAX}{}^{b}}{K_M{}^{b}}[P]}{1 + \dfrac{[S]}{K_M{}^{f}} + \dfrac{[P]}{K_M{}^{b}}}$$

where $V_{MAX}{}^{f}$ and $V_{MAX}{}^{b}$ are the maximal rates of reaction in the forward and the backward directions, respectively, and $K_M{}^{f}$ and $K_M{}^{b}$ are the Michaelis constants for the forward and the backward reactions, respectively. State explicitly how $V_{MAX}{}^{f}$, $V_{MAX}{}^{b}$, $K_M{}^{f}$ and $K_M{}^{b}$ depend on the values of rate constants k_1, k_2, k_3 and k_4.

[~30% marks]

$$V = d[P]/dt = k_3[ES] - k_4[E]_0[P] + k_4[ES][P]$$

i.e. $V = k_3.[E]_0\{k_1[S] + k_4[P]\}/(k_1[S] + k_2 + k_3 + k_4[P]) - k_4[E]_0[P]$
$+ k_4[P].[E]_0\{k_1[S] + k_4[P]\}/(k_1[S] + k_2 + k_3 + k_4[P])$

i.e. $V =$
$[E]_0\{ k_1k_3[S] + k_3k_4[P] + k_1k_4[S][P] + k_4{}^2[P]^2 - k_4[P].(k_1[S] + k_2 + k_3 + k_4[P]) \}$
$/\{k_1[S] + k_2 + k_3 + k_4[P]\}$

i.e. $V = [E]_0\{ k_1k_3[S] - k_2k_4[P] \} / \{ k_1[S] + k_2 + k_3 + k_4[P] \}$

i.e. $V = [E]_0\{ k_1k_3[S]/(k_2 + k_3) - k_2k_4[P]/(k_2 + k_3) \}$
$/ \{ k_1[S]/(k_2 + k_3) + 1 + k_4[P]/(k_2 + k_3) \}$

i.e. $V = \{ (k_3[E]_0).(k_1/(k_2 + k_3)).[S] - (k_2[E]_0).(k_4/(k_2 + k_3)).[P] \}$
$/ \{ 1 + k_1[S]/(k_2 + k_3) + k_4[P]/(k_2 + k_3) \}$

i.e. $V = \{ (V_{MAX}{}^{f}/K_M{}^{f}).[S] - (V_{MAX}{}^{b}/K_M{}^{b}).[P] \} / \{ 1 + [S]/K_M{}^{f} + [P]/K_M{}^{b} \}$

where $V_{MAX}{}^{f} = k_3[E]_0$ $\quad$ $V_{MAX}{}^{b} = k_2[E]_0$

$K_M{}^{f} = (k_2 + k_3)/k_1$ $\quad$ $K_M{}^{b} = (k_2 + k_3)/k_4$

(g) Discuss the effect on V of conducting experiments under initial rate conditions when none of the product is present.

[~10% marks]

When [P] = 0,

$$V = \{ (V_{MAX}^f/K_M^f).[S] - (V_{MAX}^b/K_M^b).0 \} / \{ 1 + [S]/K_M^f + 0/K_M^b \}$$

i.e. $V = (V_{MAX}^f/K_M^f).[S] / (1 + [S]/K_M^f)$

i.e. $V = V_{MAX}^f.[S] / (K_M^f + [S])$

Therefore under initial rate conditions when [P] = 0, the rate equation reduces to the expected form of the Michaelis-Menten equation for an ***irreversible*** reaction.

(h) Show that if this reversible enzyme catalysed reaction is allowed to run until it reaches equilibrium, the equilibrium constant, K_{eq}, will be equal to the ratio of the catalytic efficiencies of the forward and the backward reactions.

[~20% marks]

If the reversible reaction is allowed to reach equilibrium, then

$$d[P]/dt = k_3[ES] - k_4[E][P] = 0$$

i.e. $[P] = (k_3/k_4).([ES]/[E])$

Also, $d[S]/dt = k_2[ES] - k_1[E][S] = 0$

i.e. $[S] = (k_2/k_1).([ES]/[E])$

But $K_{eq} = [P]/[S]$

i.e. $K_{eq} = \{ (k_3/k_4).([ES]/[E]) \} / \{ (k_2/k_1).([ES]/[E]) \}$

i.e. $K_{eq} = (k_3/k_4) / (k_2/k_1) = (k_1k_3)/(k_2k_4)$

For the forward reaction, catalytic efficiency $= k_3/K_M^f = (k_1k_3)/(k_2 + k_3)$

For the backward reaction, catalytic efficiency $= k_2/K_M^b = (k_2k_4)/(k_2 + k_3)$

Ratio of catalytic efficiencies $= (k_3/K_M^f) / (k_2/K_M^b) = \{(k_1k_3)/(k_2 + k_3)\} / \{(k_2k_4)/(k_2 + k_3)\}$

i.e. ratio of catalytic efficiencies $= (k_1k_3)/(k_2k_4) = K_{eq}$

(This expression for K_{eq} is called Haldane's relationship.)

MB 2021 E10 Solution

(a) At equilibrium, $dN/dt = 0$. Hence,

$$r N_{eq}\left(1-\frac{N_{eq}}{k}\right)-qE N_{eq}=0$$

$$\rightarrow r\left(1-\frac{N_{eq}}{k}\right)-qE=0$$ (given that we are interested in a non-zero equilibrium)

$$\rightarrow N_{eq}=k\left(1-\frac{qE}{r}\right)$$

Equilibrium catch rate then given by

$$C_{eq}=qE N_{eq}=qkE\left(1-\frac{qE}{r}\right)$$

(b) Optimal catching effort E_{opt} maximises catch rate, which leads to the first-order maximisation condition

$$\frac{dC_{eq}}{dE}=qk\left(1-\frac{2qE}{r}\right)=0 \text{ for } E = E_{\text{opt}}$$

$$\rightarrow E_{opt}=\frac{r}{2q}$$

(c) Now

$$r N_{eq}\left(1-\frac{N_{eq}}{k}\right)-q(E_1+E_2)N_{eq}=0$$

$$\rightarrow N_{eq}=k\left(1-\frac{q(E_1+E_2)}{r}\right)$$

Equilibrium catch rate for first authority then given by

$$C_{1eq}=qE_1N_{eq}=qkE_1\left(1-\frac{q(E_1+E_2)}{r}\right)$$

Optimal catching effort for the first authority, $E_{1\text{opt}}(E_2)$, maximises catch rate, which leads to the first-order maximisation condition

$$\frac{dC_{1eq}}{dE_1}=qk\left(1-\frac{q(2E_1+E_2)}{r}\right)=0 \text{ for } E_1 = E_{1\text{opt}}(E_2)$$

$$\rightarrow E_{1opt}(E_2) = \frac{1}{2}\left(\frac{r}{q} - E_2\right)$$

The first authority's optimal catching effort decreases linearly with the second authority's effort.

(d) Stable catching effort E^{*} satisfies $E_{1opt}(E^{*}) = E^{*}$, hence

$$\frac{1}{2}\left(\frac{r}{q} - E^{\text{¿}}\right) = E^{\text{¿}}$$

$$\rightarrow E^{\text{¿}} = \frac{r}{3q}$$