HW 1

Philipp Ross

2017-01-12

Last updated: 2017-01-17

Code version: 26ce535

1. Statistics and Differential Equations Review

Review Problem 1

(a)

This is the same as a successful trail of a Bernoulli random variable, ten times in a row with probability $p = 0.5$.

```
p <- 0.5
p^10
```

```
[1] 0.0009765625
```

(b)

```
p <- 0.63095735
p^10
```

```
[1] 0.01
```

(c)

If $X \sim Binom(n, p)$, then $E[X] = np$.

Review Problem 2

(a)

```
ppois(0, 1.5, lower.tail = T)
```

```
[1] 0.2231302
```

(b)

```
ppois(1, 1.5, lower.tail = T)
```

```
[1] 0.5578254
```

(c)

The expectation of a Poisson random variable is the rate paramter, $\lambda = 1.5$. The standard deviation is the square root of the rate parameter, $\sqrt{\lambda} = 1.224745$.

Review Problem 3

(a)

$R = \frac{n(t+1)}{n(t)}$

(b)

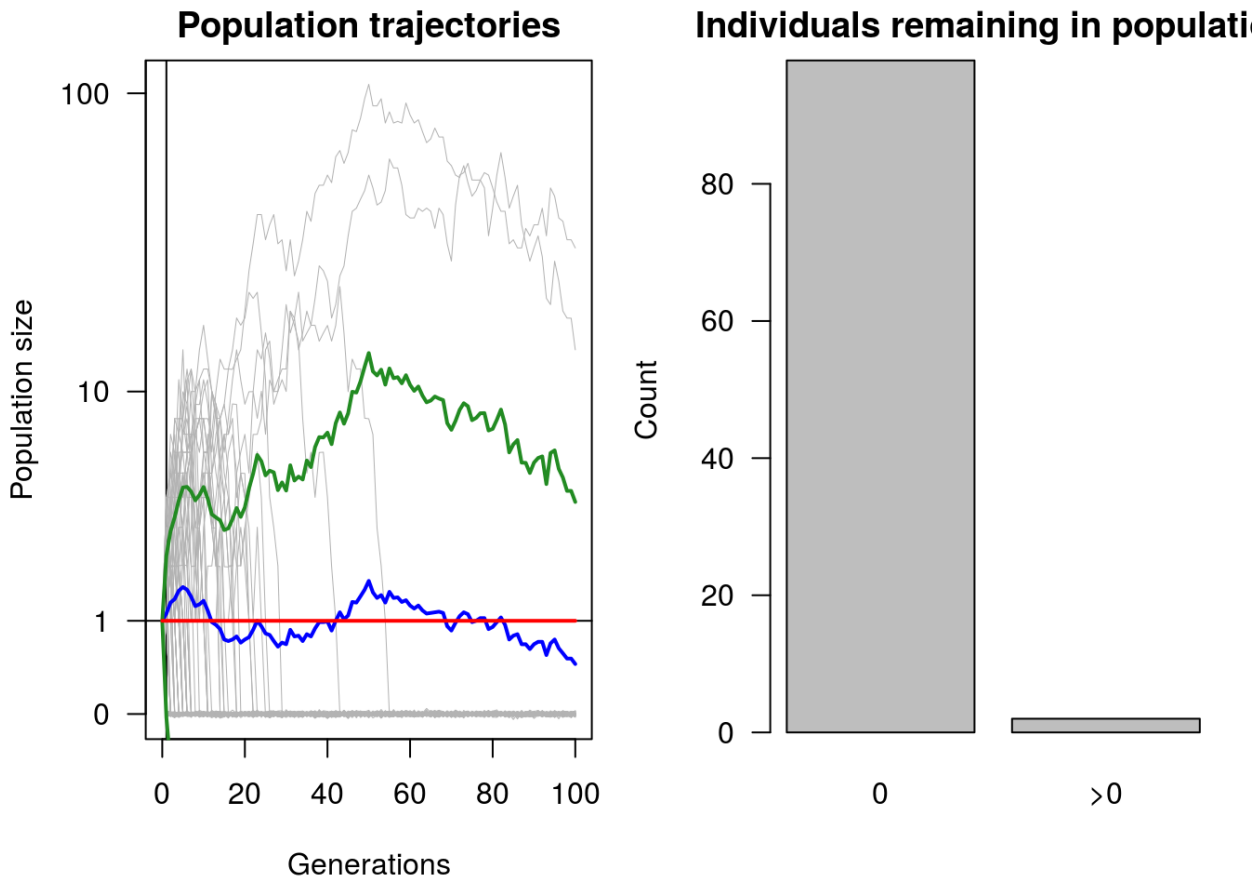
$R = \frac{n(1)}{n(0)} = \frac{n(0)e^{r \cdot 1}}{n(0)} = e^r$

2. Evolution of Antibiotic Resistance

Problem 1

(a)

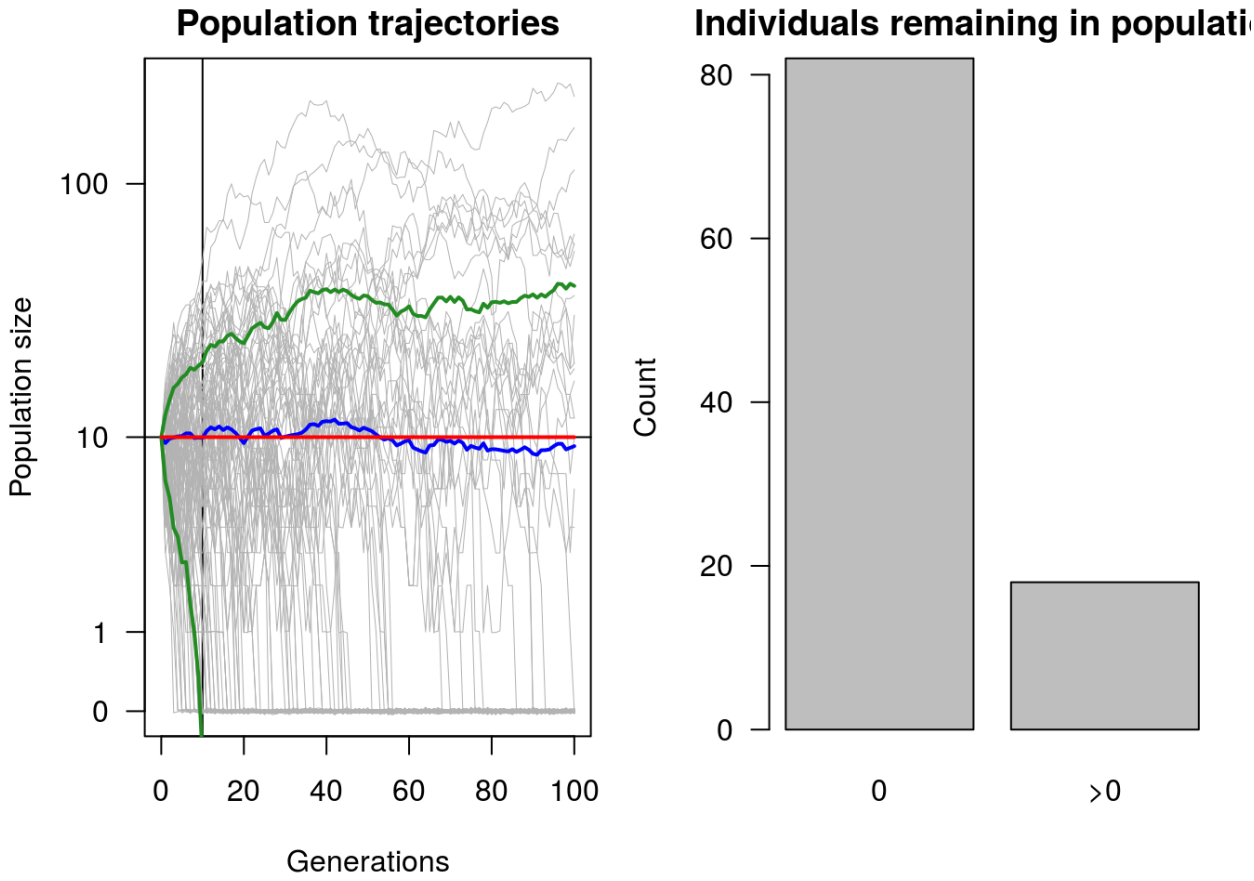
```
simExpGrowth(1,100)
```




```
Fraction of dead populations after 100 generations = 0.99  when starting with  1 individuals
```

```
simExpGrowth(10,100)
```

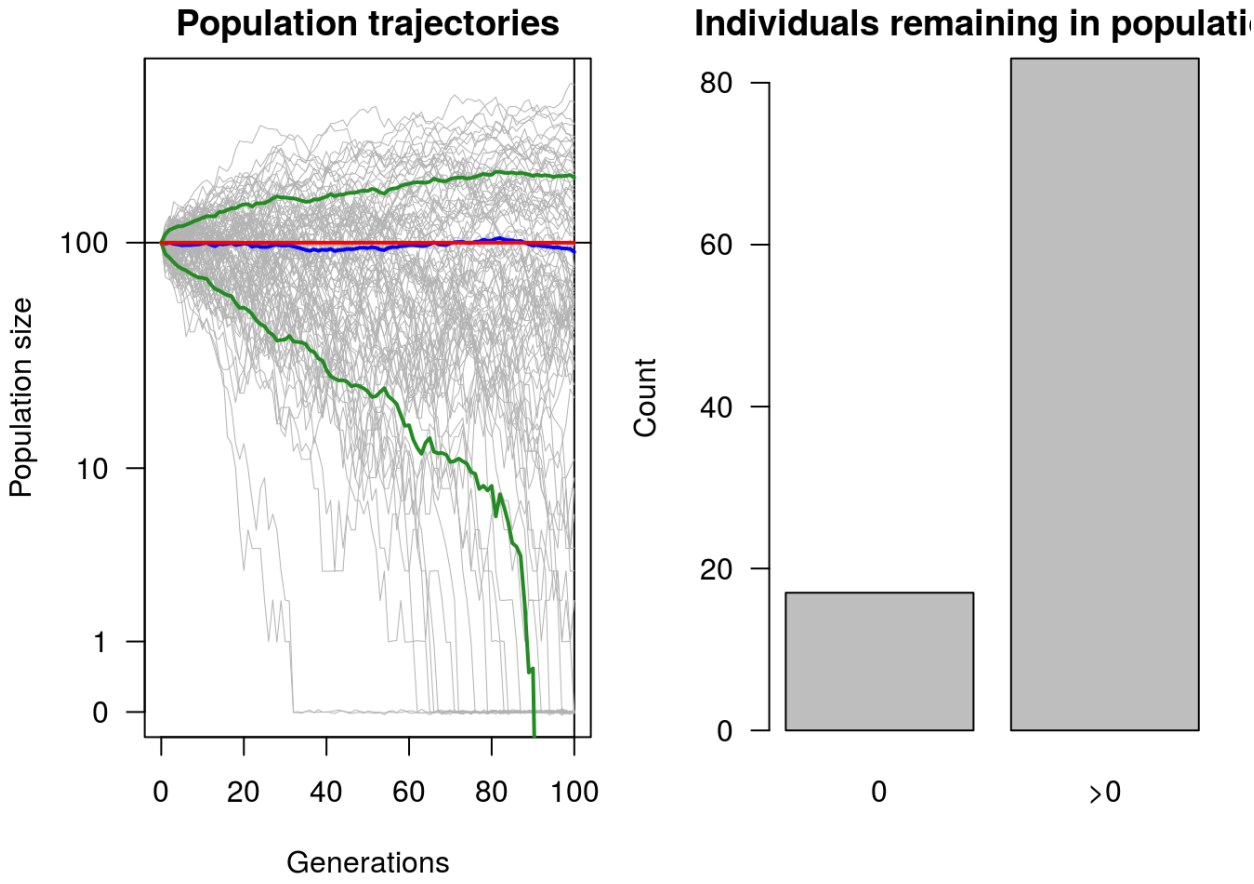
```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```




```
Fraction of dead populations after 100 generations = 0.8  when starting with  10 individuals
```

```
simExpGrowth(100,100)
```

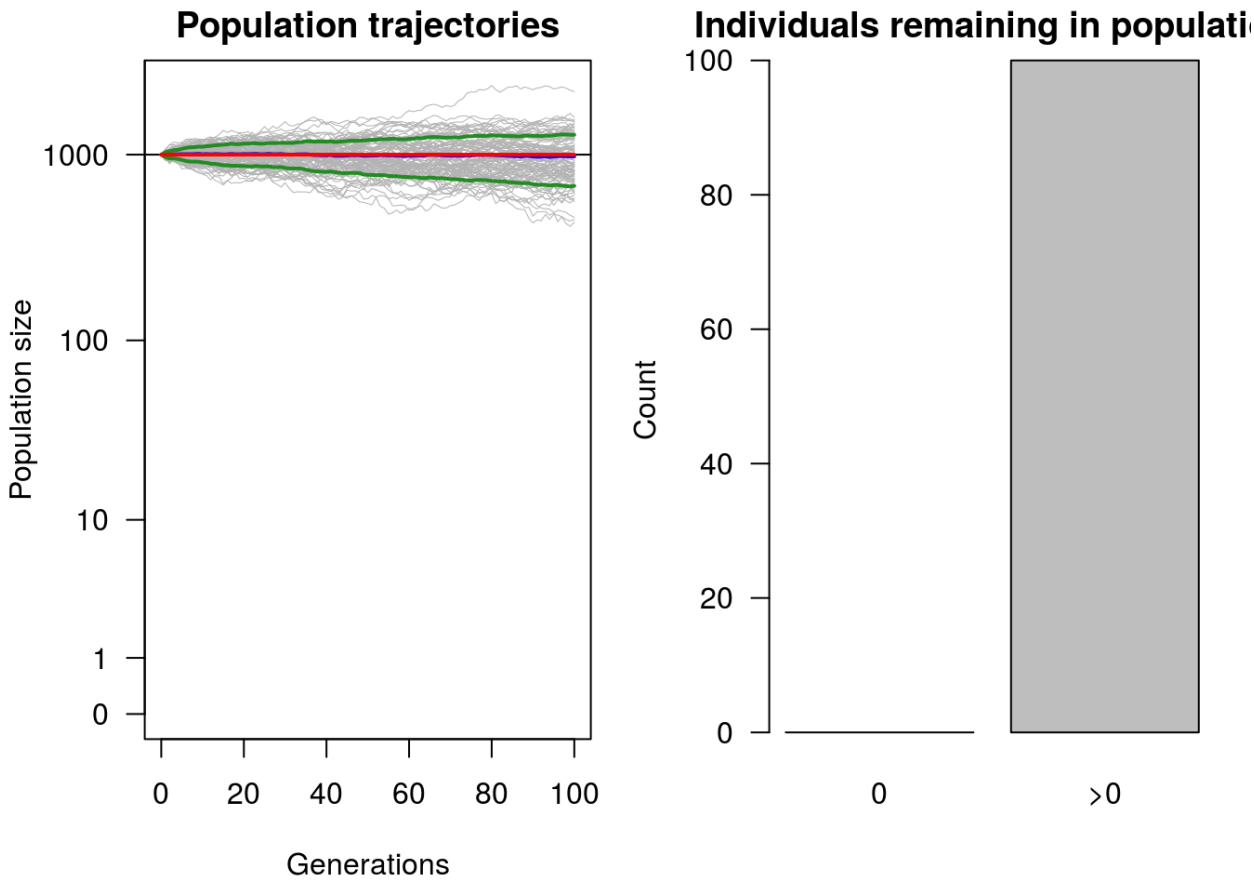
```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```




```
Fraction of dead populations after 100 generations = 0.23  when starting with  100 individuals
```

```
simExpGrowth(1000,100)
```

```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```

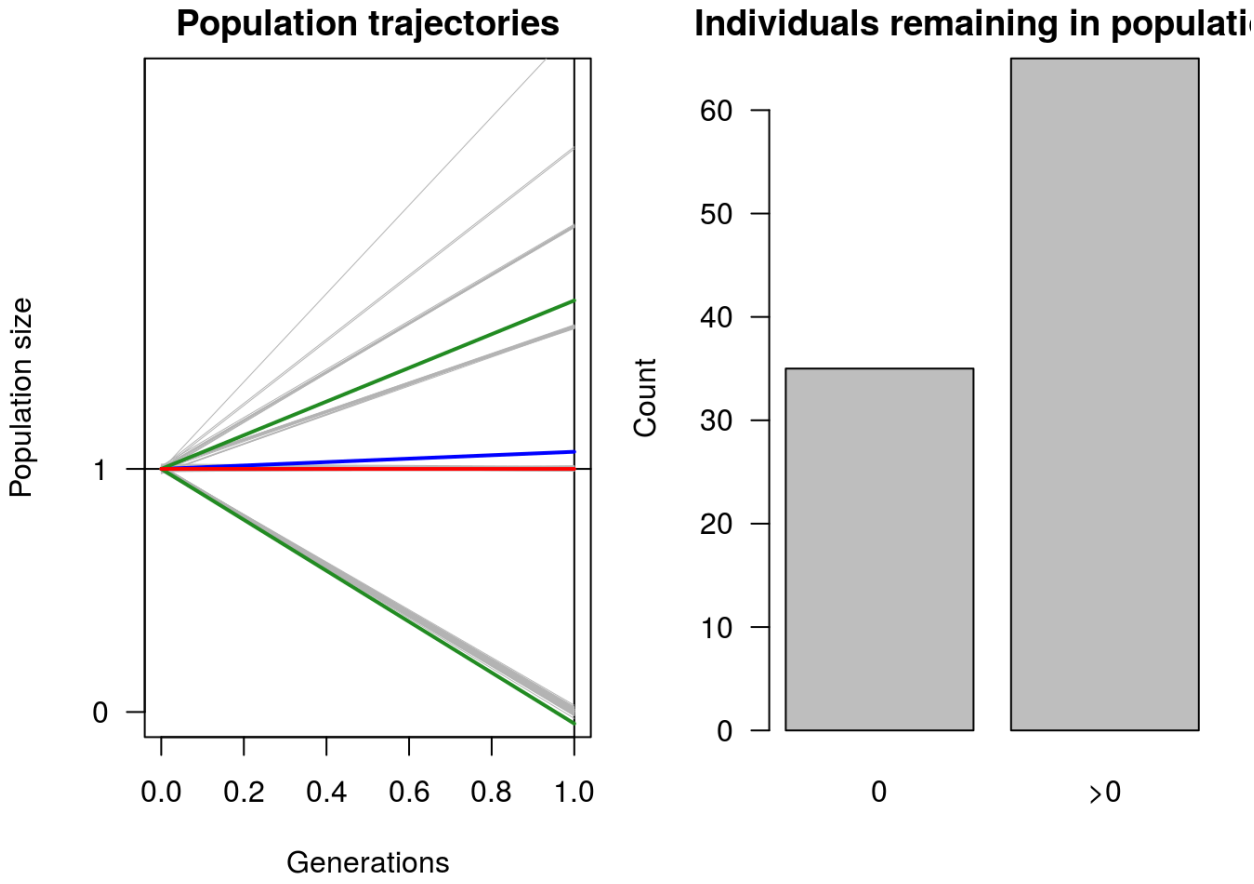



```
Fraction of dead populations after 100 generations = 0  when starting with  1000 individuals
```

(b)

```
simExpGrowth(1,1)
```

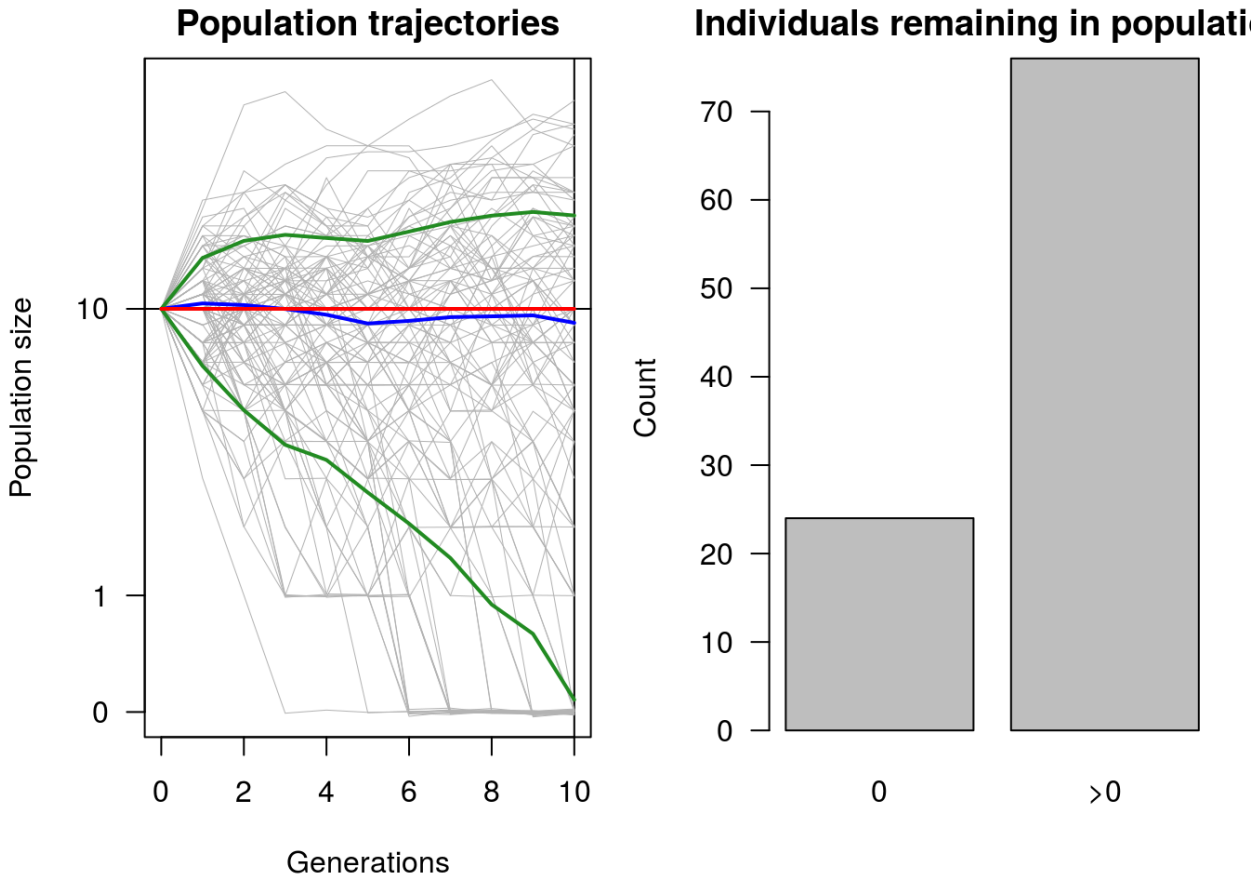
```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```




```
Fraction of dead populations after 1 generations = 0.41  when starting with  1 individuals
```

```
simExpGrowth(10,10)
```

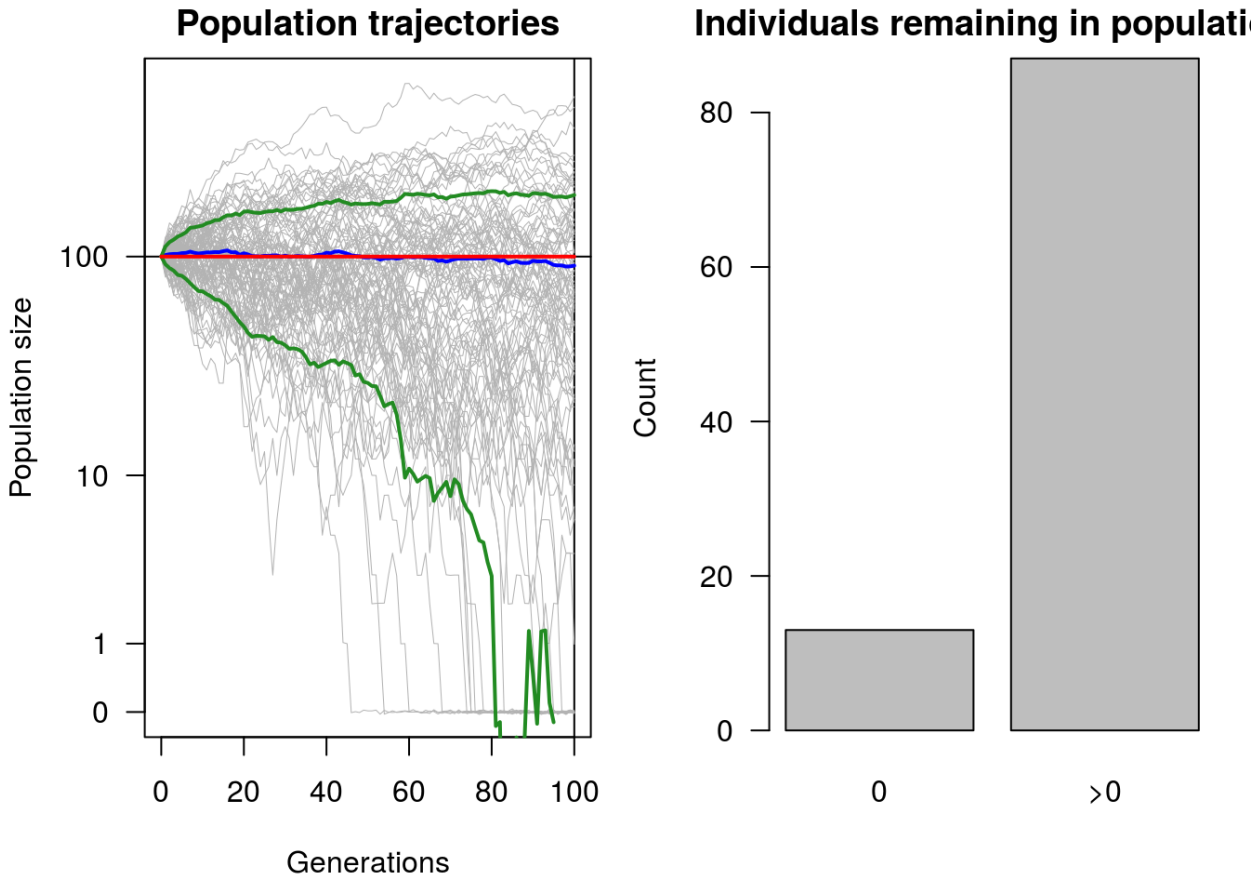
```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```




```
Fraction of dead populations after 10 generations = 0.22  when starting with  10 individuals
```

```
simExpGrowth(100,100)
```

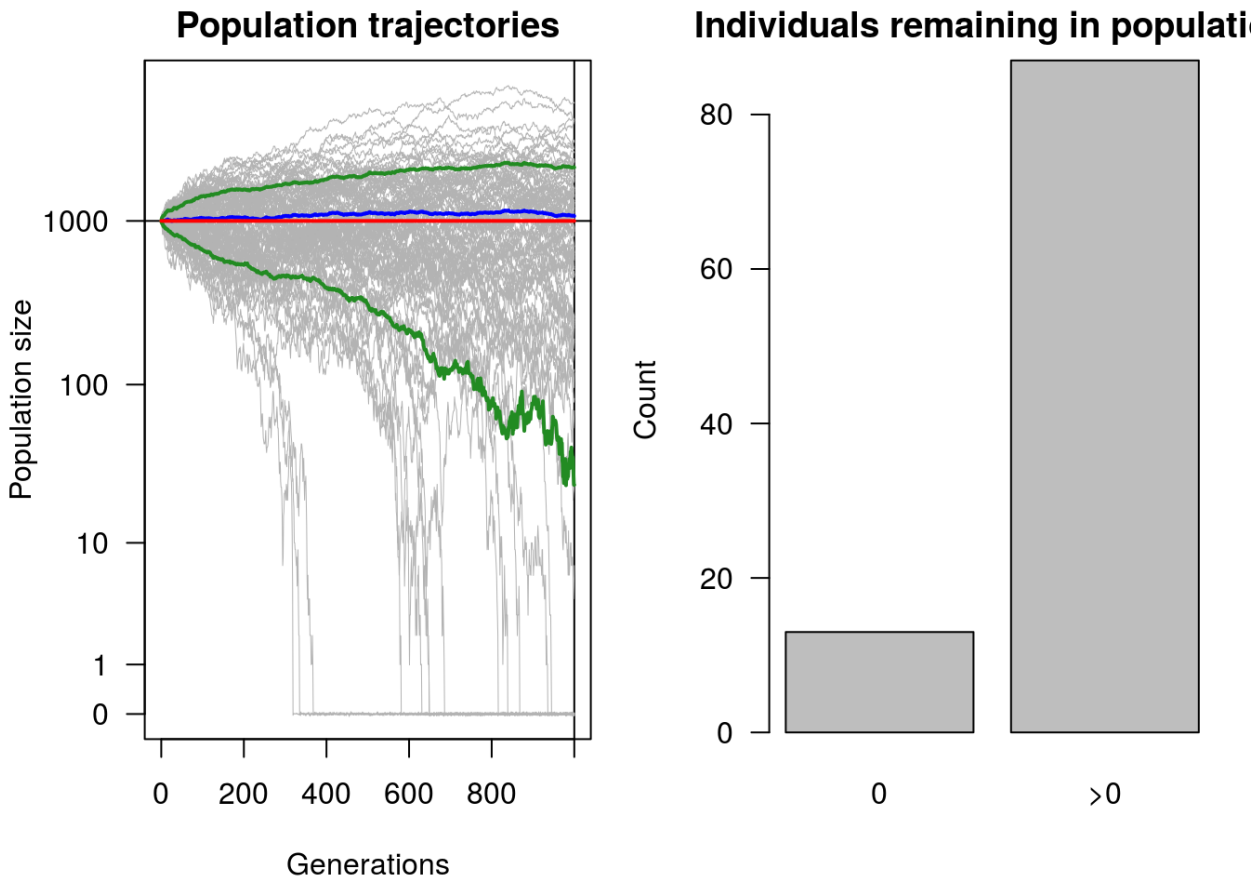
```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```




```
Fraction of dead populations after 100 generations = 0.11  when starting with  100 individuals
```

```
simExpGrowth(1000,1000)
```

```
Warning in par(new = TRUE): calling par(new=TRUE) with no plot
```



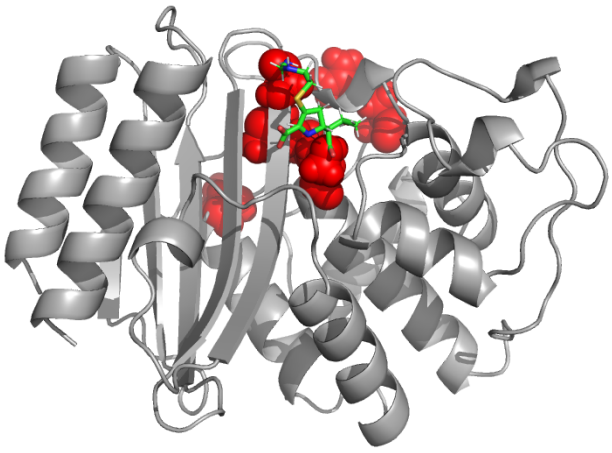

```
Fraction of dead populations after 1000 generations = 0.16  when starting with  1000 individuals
```

The fraction of dead populations is less sensitive to the

Problem 2

(a)

this is the structure of β-lactamase with an inhibitor substrate known as imipenem. THe spheres in red are the residues that most commonly accumulate changes during development of resistance to antibiotics.



(b)

The residues closest to the substrate are amino acids 69 and 70 and the residue furthest from the substrate is amino acid 182.

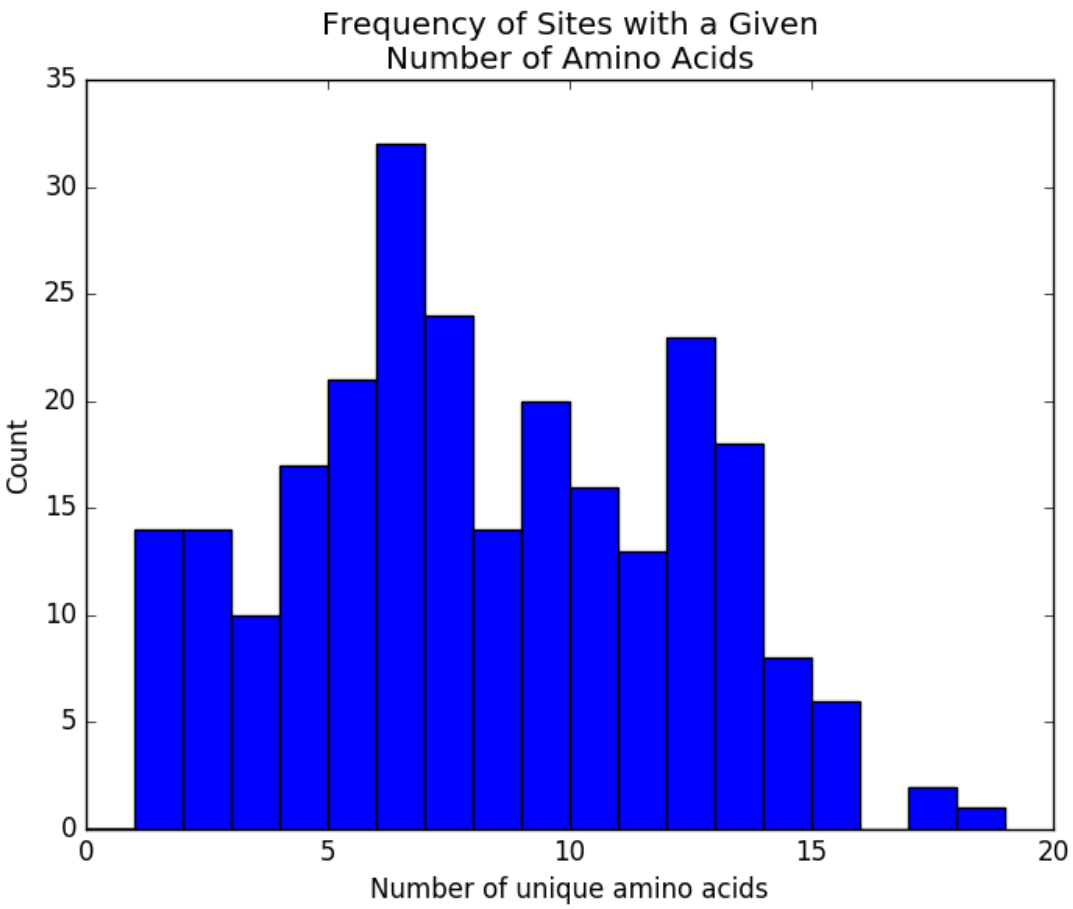
Problem 3

Once you run the script, the following is your output:

```
SingleLetterAlphabet() alignment with 125 rows and 287 columns
----------------------HPETLVKVKDAEDQLGARVGY...KHW TEM-1-imipenem-from-1BT5_PSE_File_Sequence
-------------------------LADTVQGLEERLDARIGI...--- UniRef90_Q1YJJ5_38_293/1-256
-------------------------LTQTITQIEQTLDARVGL...D-- UniRef90_A3VZZ2_27_282/1-256
---------------------------AELAELERRYGARLGV...R-- UniRef90_H2K2N0_50_303/1-254
-----------------------QGVRAELRALESRHAARLGV...R-- UniRef90_H5XLN5_39_297/1-259
-------------------------QGELRALEARHAARLGV...R-- UniRef90_H0K2N4_41_296/1-256
---------------------------RFAELETTSGARLGV...D-- UniRef90_Q9EZQ7_54_307/1-254
-------------------------AATAEFAALEQRFGARLGV...--- UniRef90_Q5YXD6_53_309/1-257
---------------------QPATLAAQLSALEASANGRLGV...AAW UniRef90_B8R7R2_34_296/1-263
-------------------------QLETPFQSLEQQHRGRLGI...Q-- UniRef90_B3PDF2_31_288/1-258
---------------------------QDRLAKLEQQLNGRLGV...EW- UniRef90_C0LNZ8_37_293/1-257
-----------------------PDAESRLATLERSAGGRLGV...--- UniRef90_Q470Y4_51_308/1-258
-------------------------AAISKLEGVEGSSGARLGV...--- UniRef90_UPI0001744437_38_284/1-247
----------------------QQDASAQLAALEAASGSRLGV...--- UniRef90_B5L5U5_32_290/1-259
---------------------------EAQLAALERSVGGRLGL...D-- UniRef90_C5CR04_40_295/1-256
-------------------------DAARQIAALESGFGGRIGV...ATF UniRef90_F2LLP9_44_303/1-260
-----------------------------------RMRGRLGV...D-- UniRef90_E5AR55_83_329/1-247
----------------------HAAAQDRLAALESAAGGRLGV...--- UniRef90_H1SON1_8_258/1-251
...
-------------------------QLNEDISLIEQQTSSRIGV...QKY UniRef90_D0MBV8_21_280/1-260
Length of aligned sequences (including gaps): 287
Number of sequences in the alignment: 125
Wrote results to AlignmentAnalysis.txt
```

```
Plotted histogram in SiteDiversityHist.png
```

(a)



Based on the plot, we can see that there around 14 invariant sites, or sites where there is but one unique amino acid within the alignment.

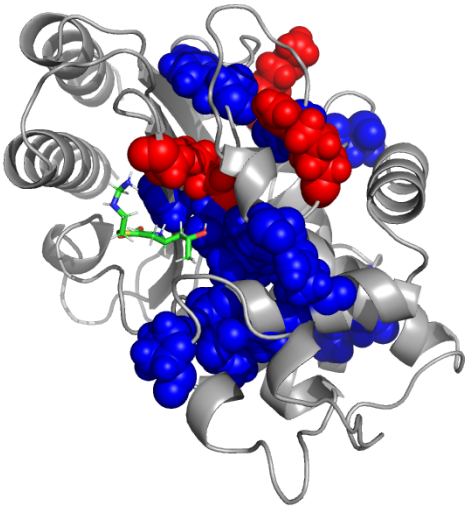
(b)

Once you modify the script, you can see that the following sites are invariant:

```
TEM-1-imipenem-from-1BT5_PSE_File_Sequence
66 F
70 S
73 K
81 L
107 P
125 A
131 D
134 A
136 N
157 D
166 E
179 D
199 L
236 G
```

(c)

The spheres highlighted in blue are the conserved amino acids while the ones in red are those relevant to the accumulation of antiobiotic resistance are in red.



(d)

The strongest noticeable relationship between the highly conserved residues and those involved in antibiotic resistance is that they only barely overlap. This suggests that those residues are highly invariant for a reason, such as that amino acid substitutions at those sites tend to be deleterious to organism fitness or have a negative consequences for protein function. It also indicates that by mutating the protein at residues that are not as highly conserved, the protein can avoid inhibition by imipenem.

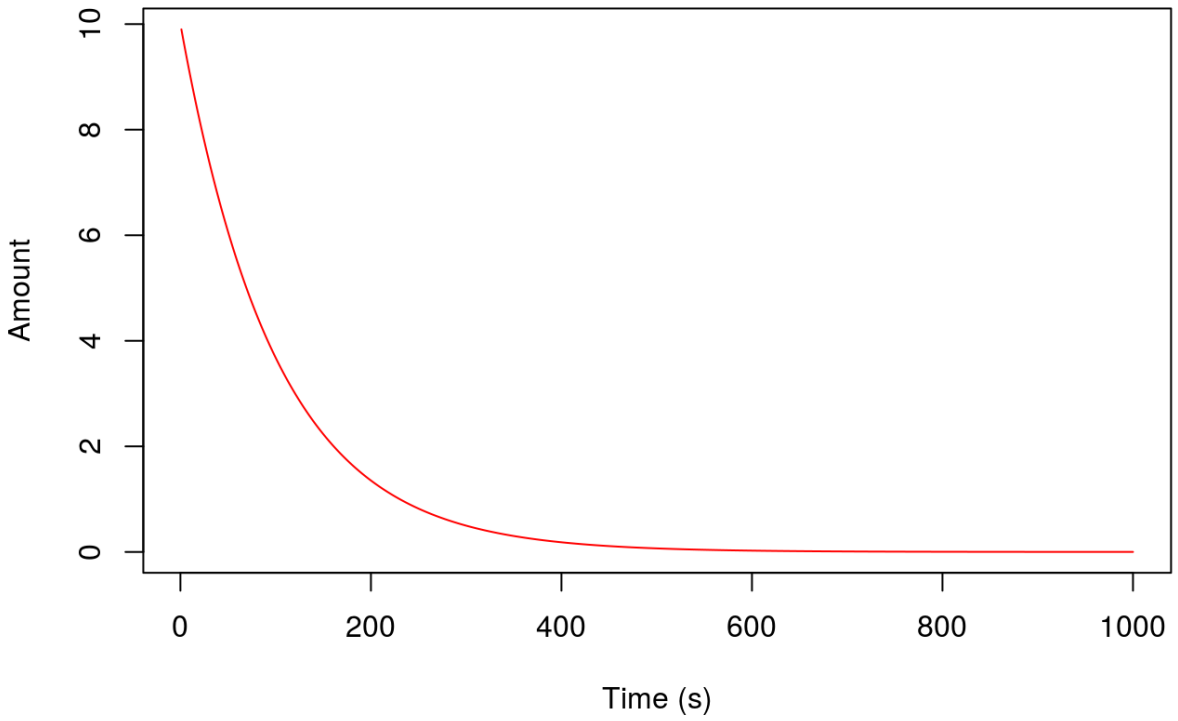
Problem 4

(a)

On average, 10μM ampicillin degrades at $0.1 \mu M s^{-1}$

(b)

```
t <- 1:1000
k <- 0.01
plot(t, 10*exp(-k*t), type="l", col="red", xlab="Time (s)", ylab="Amount")
```



(c)

k_{cat} is going to be $(10000 \cdot 0.01) - 0.01 = 99.99$.

(d)

k_m is going to be $\frac{99.99+10}{1000} = 0.10999$

(e)

Now that the initial concentration include β-lactamase, the average degradation rate will be much faster. It will be:

$$(9 \cdot 0.01) + (1 \cdot 100) = 100.09 \mu M s^{-1}$$

Session Information

```
sessionInfo()
```

```
R version 3.3.1 (2016-06-21)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 16.04.1 LTS

locale:
```

```
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

loaded via a namespace (and not attached):
 [1] backports_1.0.4 magrittr_1.5    rprojroot_1.1   tools_3.3.1
 [5] htmltools_0.3.5 yaml_2.1.14     Rcpp_0.12.6     stringi_1.1.1
 [9] rmarkdown_1.3   knitr_1.15.1    git2r_0.18.0    stringr_1.1.0
[13] digest_0.6.11   workflowr_0.2.0 evaluate_0.10
```