DEMG609: Problem Set 1

Nick Graetz

September 13, 2017

1.(a)

```r
# Because the instantaneous growth rate is constant over the interval, it can be calculated as the mean annualized
N_t1 <- 281421906
N_t2 <- 308745538
t1 <- 2000.25
t2 <- 2010.25
T <- t2 - t1
r_star <- log(N_t2 / N_t1) / T
answer_1a <- r_star
```

Answer: 0.0092662

1.(b)

```r
tripling_time <- log(3) / r_star
answer_1b <- tripling_time
```

Answer: 118.5606806

1.(c)

```r
# Given our growth rate is constant...
py <- ((N_t2 - N_t1) * T) / log(N_t2 / N_t1)
answer_1c <- py
```

Answer: 2948727626.35409

1.(d)

```r
# What was the population size in the middle of the intercensual period?
# Given that growth rate is constant...
pop_at_t <- function(t) {
  new_pop <- N_t1 * exp(r_star * t)
  return(new_pop)
}
N_mid <- pop_at_t(T/2)
answer_1d <- N_mid
```

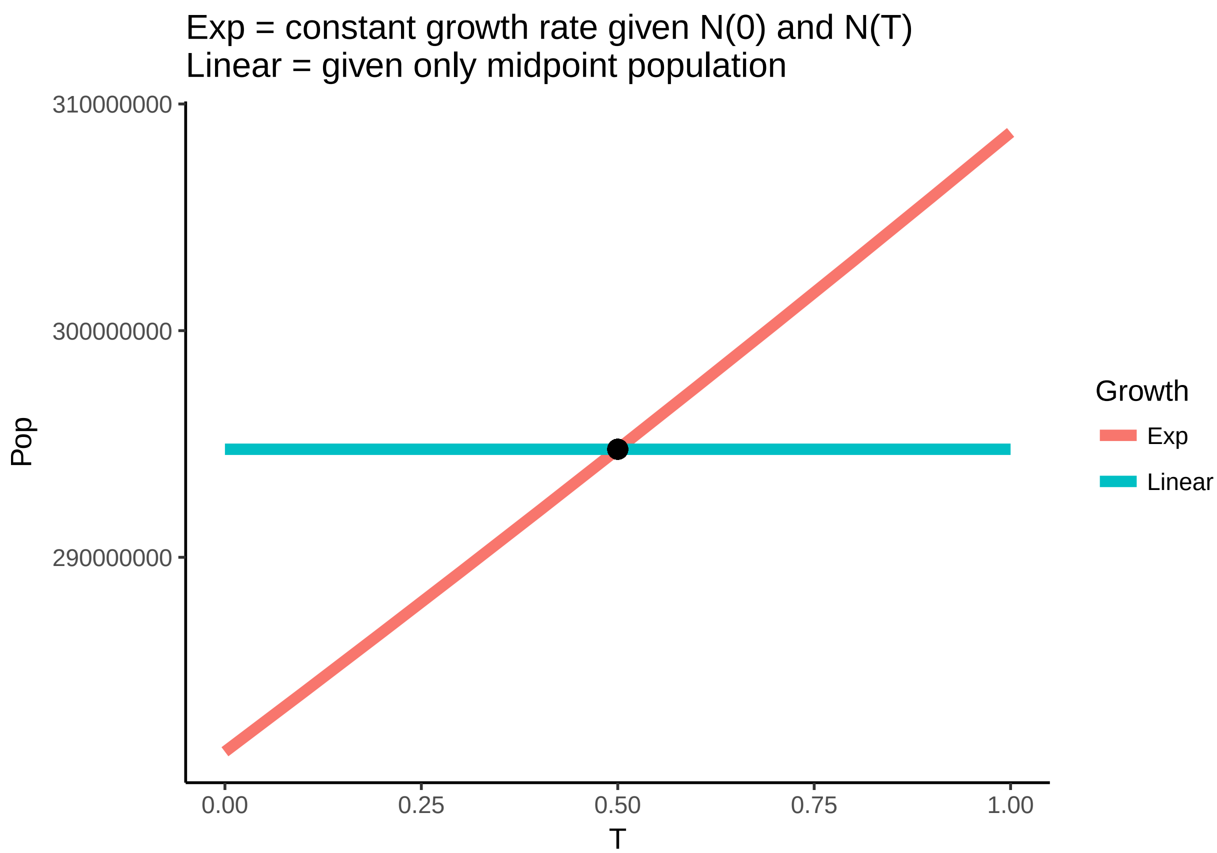
Answer: 294767294.272881

1.(e)

```r
# Given only midpoint population, we have to assume linear growth
py_linear <- N_mid * T

growth <- data.table(T = seq(0,1,0.1),
                     Pop = c(rep(N_mid, 11),
                             pop_at_t(0:10)),
                     Growth = c(rep('Linear', 11),
                                rep('Exp', 11)))

ggplot(data = growth) +
  geom_line(aes(x = T,
                y = Pop,
                color = Growth),
            size = 2) +
  geom_point(x = 0.5, y = N_mid, color = 'black', size = 3) +
  ggtitle("Exp = constant growth rate given N(0) and N(T)\nLinear = given only midpoint population") +
  theme_classic()
```



Answer:
Linear PY: 2947672942.72881
Exp PY: 2948727626.35409
Difference (Linear - Exp): -1054683.6252751

When N_t follow an exponential growth pattern, $N_{\frac{T}{2}} * T$ will underestimate person-years lived during the period.

2.(a)

Answer: 0.0092662

Even though the instantaneous growth rate varies over the period, we must estimate it as the mean annualized growth rate over the period and assume the instantaneous rate is constant. This is a logistical assumption, as we can never measure the actual instantaneous growth rate continuously over the entire period.

3.(a)

```
# Instead of 2000.25 to 2010.25, it's now 2000.25 - 2010.08333
N_t1 <- 281421906
N_t2 <- 308745538
t1 <- 2000.25
t2 <- 2010 + (1/12)
T <- t2 - t1
r_star <- log(N_t2 / N_t1) / T
answer_3a <- r_star
```

Answer: 0.0094233

3.(b)

```
pop_at_t <- function(t, r_star, N_t1) {
  new_pop <- N_t1 * exp(r_star * t)
  return(new_pop)
}
N_mid <- pop_at_t(t = (T/2), r_star = r_star, N_t1 = N_t1)
answer_3b <- N_mid
# Bigger, because it has to get to the same end population faster
```

Answer: 294767294.272881
Answer to 1(d): 294767294.272881

3.(c)

Given that N_0 and N_t and the assumption of constant r between N_0 and N_t: $N_t = N_0 * e^{rt}$ and $r = ln\left(\frac{N_t}{N_0}\right)$

$$N_t = N_0 e^{\frac{ln\left(\frac{N_t}{N_0}\right)}{T} T}$$

$$N_t = N_0 e^{ln\left(\frac{N_t}{N_0}\right)}$$

At this point, we have an equation that does not contain r or T, only N_t and N_0. Below, reducing to make sure formula still internally consistent.

$$N_t = N_0 * \frac{N_t}{N_0}$$

$$N_t = N_t$$

$$1 = 1$$

4.(a)

```r
# Reset to initial pops and dates
N_t1 <- 281421906
N_t2 <- 308745538
t1 <- 2000.25
t2 <- 2010.25
T <- t2 - t1
# Inputs:
#    j = number of times to compound per year
#    year_interval = years in the period
#    start_pop
#    end_pop
# Outputs:
#    r_j = annual geometric growth rate
calc_geo_r <- function(j, year_interval, start_pop, end_pop) {
  r_j <- j * ((end_pop / start_pop)^(1/(j*year_interval)) - 1)
  return(r_j)
}

# Apply function over j = 1, 12, 365 (annually, monthly, daily, continuously)
geom_rates <- lapply(c(1,12,365,1000000),
                     calc_geo_r,
                     year_interval = T,
                     start_pop = N_t1,
                     end_pop = N_t2)
```

Answer:
Annually: 0.0093093

Monthly: 0.0092698
Daily: 0.0092664
Cont: 0.0092662

4.(b)

```
calc_geo_r_2 <- function(j, annual_rate) {
  r_j <- j * ((1 + annual_rate)^(1/j) - 1)
  return(r_j)
}
r_j_cont <- calc_geo_r_2(j = 1000000, annual_rate = geom_rates[[1]])
answer_4b <- r_j_cont
```

Answer: 0.0092662

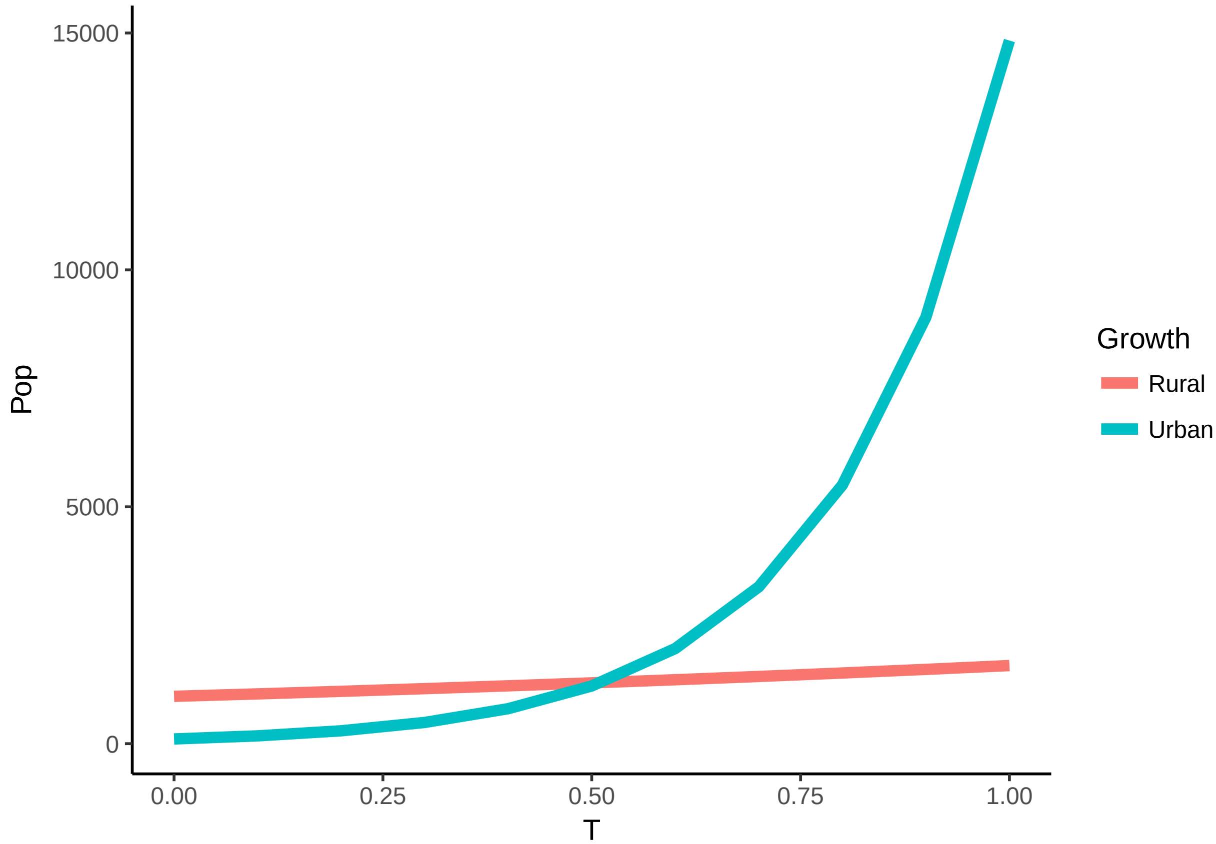
(B)

```
r_star_ukraine <- -0.49 / 100
halving_time <- log(0.5) / r_star_ukraine
answer_4B <- halving_time
```

Answer: 141.4586083

(C)

```
t1_urban <- 100
t1_rural <- 1000
r_star_urban <- 0.5
r_star_rural <- 0.05
growth_comparison <- data.table(T = seq(0,1,0.1),
                                Pop = c(pop_at_t(0:10, r_star = r_star_urban, N_t1 = t1_urban),
                                        pop_at_t(0:10, r_star = r_star_rural, N_t1 = t1_rural)),
                                Growth = c(rep('Urban', 11),
                                           rep('Rural', 11)))
ggplot() +
  geom_line(data = growth_comparison,
            aes(x = T,
                y = Pop,
                color = Growth),
            size = 2) +
  theme_classic()
```




```
growth_comparison <- dcast(growth_comparison, T ~ Growth, value.var = "Pop")
growth_comparison[, pop_ratio := Urban / Rural]
t1_ratio <- growth_comparison[T == 0, pop_ratio]
t2_ratio <- growth_comparison[T == 1, pop_ratio]
ratio_growth_rate <- log(t2_ratio / t1_ratio) / 10
ratio_growth_rate == (r_star_urban - r_star_rural)
```

```
## [1] TRUE
```

Example given my previously defined functions and arbitrary starting values:
Urban growth rate: 0.5
Rural growth rate: 0.05
Difference in growth rates: 0.45
Growth rate of urban/rural population ratio: 0.45

Analytic Answer:

$$R = \frac{N_u}{N_r}$$

Prove that: $Growth_R = Growth_u - Growth_r$

$$\frac{ln\left(\frac{R_t}{R_0}\right)}{T} = Growth_u - Growth_r$$

$$\frac{ln\left(\frac{\frac{N_{t,u}}{N_{t,r}}}{\frac{N_{0,u}}{N_{0,r}}}\right)}{T} = \frac{ln\left(\frac{N_{t,u}}{N_{0,u}}\right)}{T} - \frac{ln\left(\frac{N_{t,r}}{N_{0,r}}\right)}{T}$$

$$ln\left(\frac{\frac{N_{t,u}}{N_{t,r}}}{\frac{N_{0,u}}{N_{0,r}}}\right) = ln\left(\frac{N_{t,u}}{N_{0,u}}\right) - ln\left(\frac{N_{t,r}}{N_{0,r}}\right)$$

Use this law to reduce all terms: $ln\left(\frac{a}{b}\right) = ln(a) - ln(b)$

$$ln\left(\frac{N_{t,u}}{N_{t,r}}\right) - ln\left(\frac{N_{0,u}}{N_{0,r}}\right) = ln\left(\frac{N_{t,u}}{N_{0,u}}\right) - ln\left(\frac{N_{t,r}}{N_{0,r}}\right)$$

$$ln(N_{t,u}) - ln(N_{t,r}) - ln(N_{0,u}) - ln(N_{0,r}) = ln(N_{t,u}) - ln(N_{0,u}) - ln(N_{t,r}) - ln(N_{0,r})$$

$$1 = 1$$