DATA 609: Homework 10

Modeling with a Differential Equation

Aaron Grzasko

October 27, 2017

```
library(tidyr)
library(dplyr)
library(knitr)
library(ggplot2)
library(ggthemes)
library(car)
library(latex2exp)
```

Page 469, Question 3

The following data were obtained for the growth of a sheep population introduced into a new environment on the island of Tasmania (adapted from J. Davidson, "On the Growth of the Sheep Population in Tasmania," *Trans. R. Soc. S. Australia* 62 (1938): 342-346).

```
# data
t <- c(1814,1824,1834,1844,1854,1864)
P.t <- c(125,275,830,1200,1750,1650)
df <- data.frame(t=t,P.t=P.t)
```
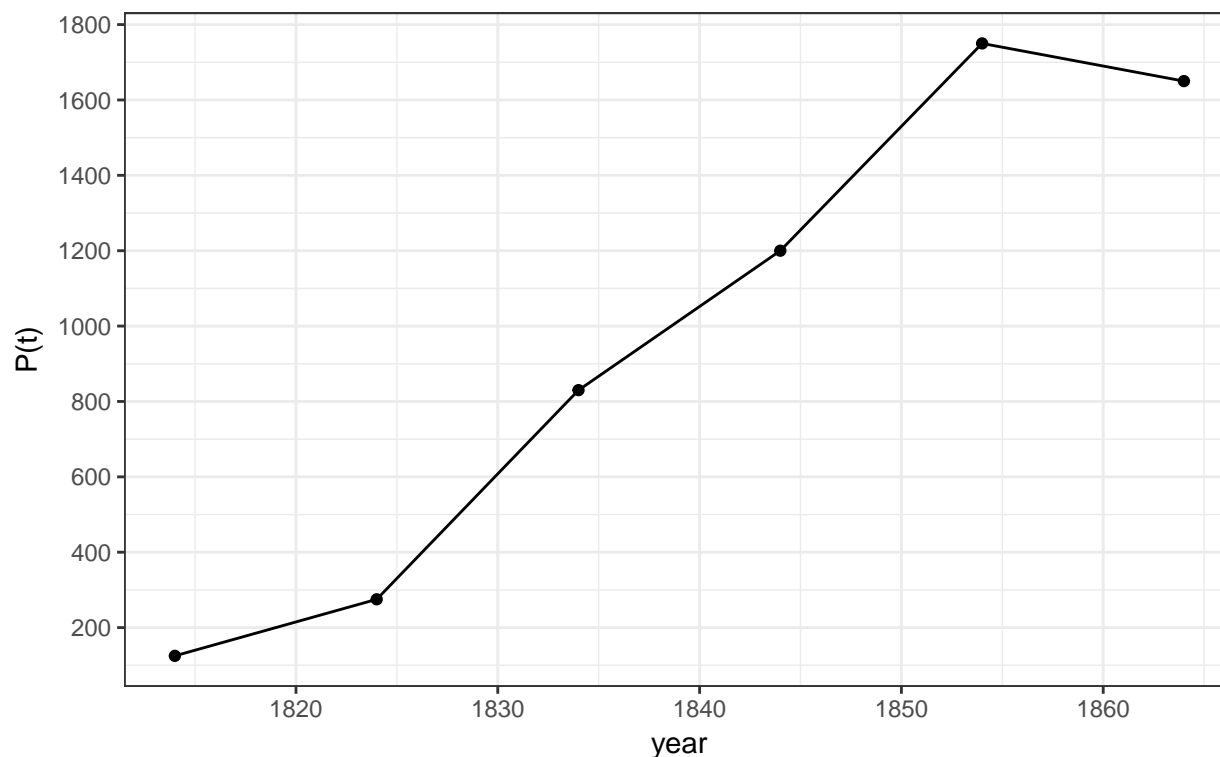
t (year)	1814	1824	1834	1844	1854	1864
P(t)	125	275	830	1200	1750	1650

Part A) Make an estimate of *M* by graphing *P(t)*.

```
ggplot(df, aes(t,P.t)) + theme_bw() +
  labs(title = "Sheep Population by Year", y = "P(t)",
       subtitle = "Tasmania (1814-1864)", x = "year") +
  scale_y_continuous(breaks = seq(0,1800,200)) + geom_point() +
  geom_line()
```

Sheep Population by Year
Tasmania (1814–1864)

If we assume a limited growth model (i.e. logistic growth curve), then the population will approach a long-term, maximum population value, M, for large values of t, time. The graph above is based on only six data points; so our visual approximation to M will likely not extrapolate well to periods beyond those in the sample, even if the logistic growth model is appropriate.

We note that the population for the periods examined reaches a maximum value of 1750 in the year 1854, and drops to 1650 in 1864. Given this limited information, we conjecture that the value 1750 is somewhat close to the limiting value of M. For now, we'll peg the value of M to be in the range of 1700 and 1900.

Part B) Plot $ln[P/(M-P)]$ against t. If a logistic curve seems reasonable, estimate rM and t^*.

Our logistic model starts with the following equation:

$$\frac{dP}{dt} = r(M - P)$$

assuming r and M are constants.

After integrating and algebraic manipulation, we have

$$ln\frac{P}{M - P} = rMt + C$$

with C referring to the constant of integration.

In the equation above, we see that $ln\frac{P}{M-P}$ is linearly related to time, t.

If the logistic curve is a reasonable model for our sheep population data, then a plot of $ln\frac{P}{M-P}$ vs. time should be approximated reasonably well by a line.

```
# midpoint of estimated range for M from part a
m <- 1800

# plot
ggplot(df, aes(t,log(P.t / (m - P.t)))) + geom_point() +
  stat_smooth(method = 'lm', se=FALSE, color = 'black') +
  labs(title = "ln[P/(M-P)] vs. Time", x = 'year',
       y = "ln[P/(M-P)]", subtitle = "with fitted OLS line") +
  theme_bw()
```


ln[P/(M−P)] vs. Time
with fitted OLS line
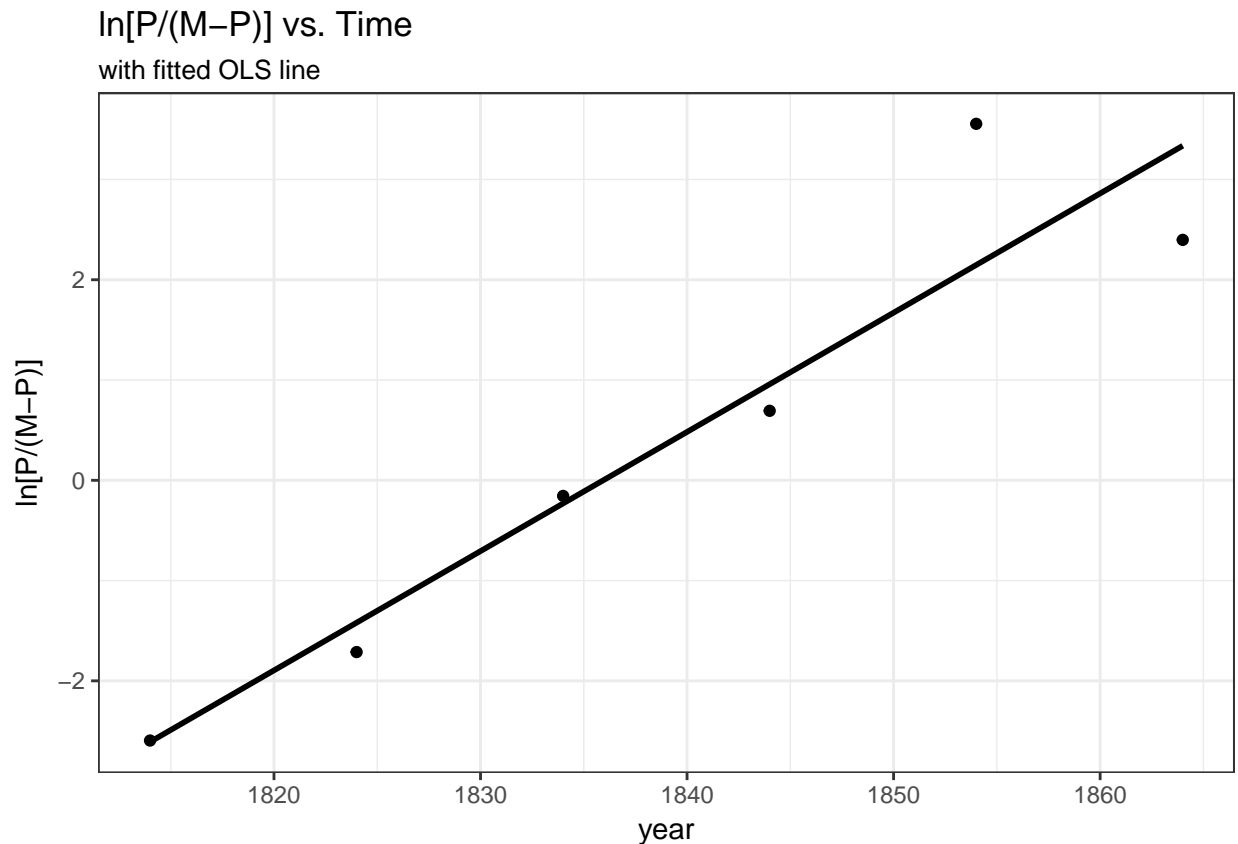
The linear model fits the data fairly well; so the logistic growth model is a reasonable approximation..

The value rM is the slope of the our fitted ols line:

```
# calculate rM
rM <- coef(lm(log(P.t/(m-P.t))~t, data=df))[2]
names(rM) <- NULL
rM
```

```
[1] 0.1189136
```

```
# given M = 1800, calculate r
r <- rM / m
r
```

```
[1] 6.606309e-05
```

Now we'll calculate t^*, which denotes the time when the population P reaches half of the limiting value.

$$t^* = t_0 - \frac{1}{rM} ln \frac{P_0}{M - P_0}$$

```r
# t0, P0 from sheep df
t0 <- df[1,"t"]
P0 <- df[1,"P.t"]

# calculate t*
t_star <- t0 - 1 / rM * log(P0 / (m-P0))
t_star
```

```
[1] 1835.825
```

Finally, let's fit a logistic model to the sheep data using R functions and plot the result.

```r
# initial estimates of model parameters phi2, phi3
phi2_initial <- coef(lm(logit(P.t * 1/1750)~t,data=df))[1]
phi3_initial <- coef(lm(logit(P.t * 1/1750)~t,data=df))[2]

# fit model, use initial estimate for M, i.e. 1800 ,from previous work
logistic <- nls(P.t~phi1/(1+exp(-(phi2+phi3*t))),
 start=list(phi1=m,phi2=phi2_initial,phi3=phi3_initial),data=df)

# logistic function with fitted, parameter estimates
phi1 <- coef(logistic)[1]
phi2 <- coef(logistic)[2]
phi3 <- coef(logistic)[3]
func <- function(t) phi1 / (1+exp(-(phi2 + phi3 * t)))

# plot fitted logistic function
ggplot(df, aes(t,P.t)) + theme_bw() +
  labs(title = "Sheep Population by Year", y = "P(t)",
       subtitle = "with fitted logistic growth curve", x = "year") +
  scale_y_continuous(breaks = seq(0,1800,200)) + geom_point() +
  stat_function(fun = func) + xlim(1814,1900 )
```

Sheep Population by Year
with fitted logistic growth curve
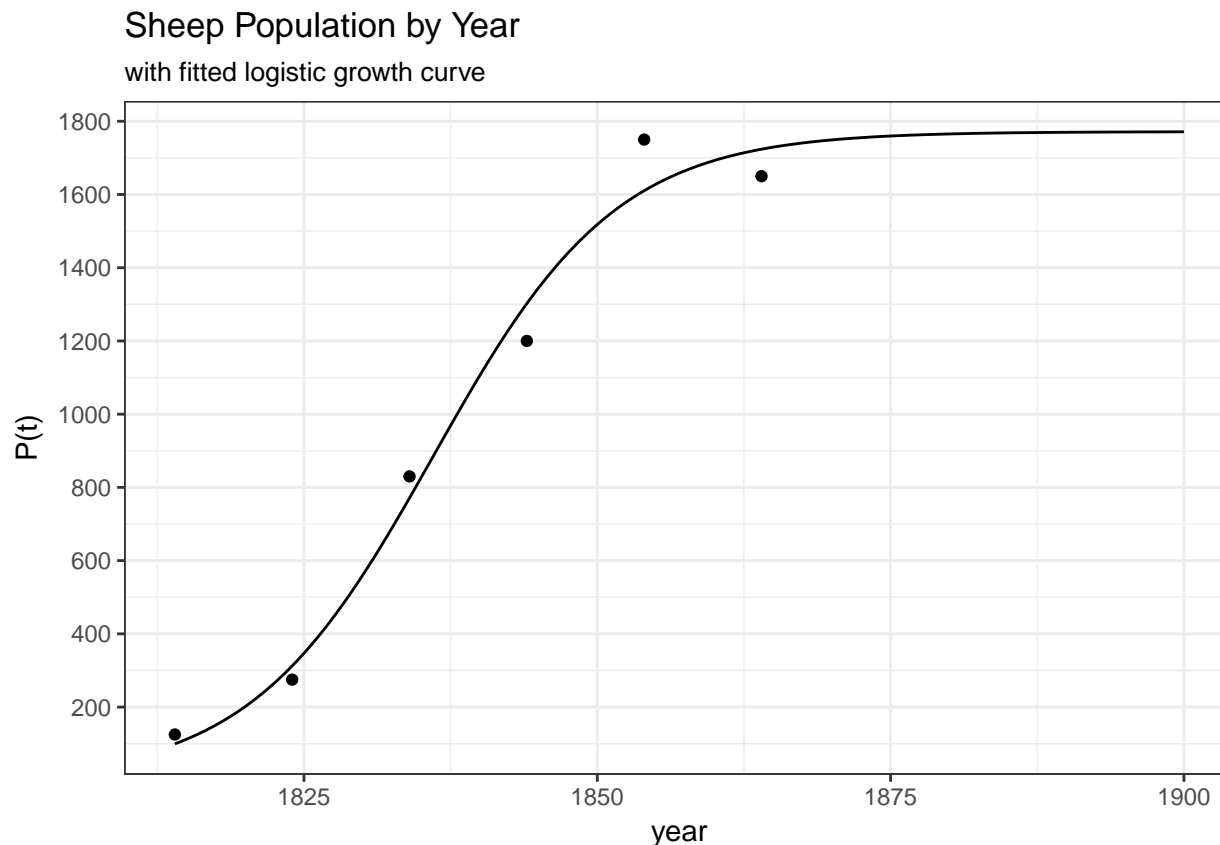
From our logistic model above, we see that the estimated maximum value for the population is 1772.

Page 478, Question 6

Suggest other phenomena for which the Prescribing Drug Dosage model described in the text might be used.

Suppose that you want to establish a no-interest checking account for your son, an impulsive college student. You want the account to maintain a value between \$L and \$H in the long run. Your son will frequently withdraw funds from this account–assume a continuous approximation is reasonable–in amounts that are proportional to total funds available at the time of withdrawal. You want to deposit a flat amount into the account at regular time intervals. Using the drug dosage model, you can solve for both of these quantities.

Alternatively, you may wish to initially fund the account so the account's value is \$H. can use the drug dosage model to determine appropriate subsequent flat deposits and regular time intervals so that the value of the account maintains a value between \$L and \$H.

Page 481, Question 1

Part A) Using the estimate that $d_b = 0.054v^2$, where 0.054 has dimension $ft - hr^2/mi^2$, show that the constant k in the following equation has the value 19.9 ft/sec^2.

The relevant equation is

$$d_b = \frac{-v_0^2}{2k} + \frac{v_0^2}{k} = \frac{v_0^2}{2k}$$

$$k = \frac{1}{2d_b}$$

```r
# seconds per hr
sec_hr <- 60^2

# ft per mile
ft_mile <- 5280

# convert 0.054 from ft-hr^2 / mi^2 to sec^2 / ft
conv_d <- 0.054 * sec_hr^2 / ft_mile^2

# calculate k in ft/sec^2
k <- 1 / (2* conv_d)
k
```

```
[1] 19.9177
```

Part B) Using the data in Table 4.4, plot d_b in ft versus $v^2/2$ in ft^2/sec^2 to estimate $1/k$ directly

Note: We assume the authors intended to reference to Table 2.4 rather than Table 4.4, as Table 2.4 does not include braking distances.

speed (mph)	20	25	30	35	40	45	50	55	60	65	70	75	80
braking dist (ft)	20	28	40.5	52.5	72	92.5	118	148.5	182	220.5	266	318	376

```r
# speed mph and breaking dist ft
speed <- c(20,25,30,35,40,45,50,55,60,65,70,75,80)
db <- c(20,28,40.5,52.5,72,92.5,118,148.5,182,220.5,266,318,376)

# speed, converted to ft/sec from miles per hour
speed_ft.sec <- speed * 5280/60^2

# speed ft/sec squared, divided by 2
speed_sq_div2 <- (speed_ft.sec^2)/2

# save to df
df <- data.frame(speed_sq_div2=speed_sq_div2, braking_dist = db)

# fit linear, proportionality model bd vs speed ^2 / 2
mylm <- lm(braking_dist~speed_sq_div2 + 0,data=df)
myfunc = function(x) coef(mylm)*x    # function for proportionality model

# plot braking distance vs v^2 / 2
ggplot(df, aes(speed_sq_div2,braking_dist)) + geom_point() +
  theme_bw() + labs(x=TeX('$v^2/2$ $(ft^2/sec^2)$'), y=TeX('$d_b$ (ft)'),
                title=TeX("Braking Distance vs. $v^2/2$"),
                        subtitle=TeX("with $y =\\beta x$ best-fit line")) +
  stat_function(fun = myfunc)
```
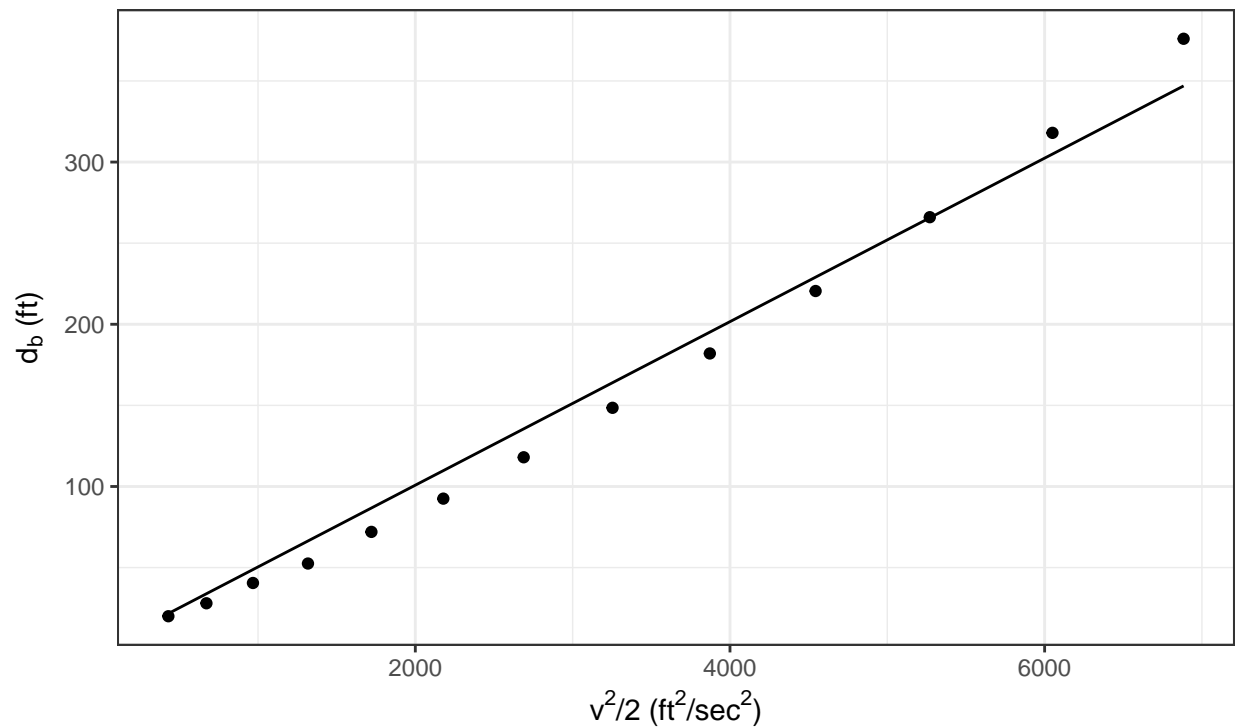
Braking Distance vs. $v^2/2$

with $y = \beta x$ best–fit line


```r
# k is inverse of slope of OLS line
k <- 1/coef(mylm)
names(k) <- NULL
k
```

```
[1] 19.84103
```

References

- logistic growth curve in R: https://bscheng.com/2014/05/07/modeling-logistic-growth-data-in-r/