Classwork 12: Log-linear and Log-log Models

Some hints for this assignment:

Log rules:
Definition: If $\log(a) = b$, then $e^b = a$.
$\log(ab) = \log(a) + \log(b)$
$\log(\frac{a}{b}) = \log(a) - \log(b)$
$\log(a^b) = b\log(a)$

In R, you can calculate (natural) logs using `log()`.
You can calculate e^a using `exp(a)`.

Part 1: Log-linear Models

A log-linear model has $log(y)$ on the left and linear terms on the right: $log(y_i) = \beta_0 + \beta_1 x_i + u_i$

What kind of process yields this kind of relationship? Consider the formula for exponential growth or decay:

$$y = (initial\ amount)\ e^{rt} \tag{1}$$

1) Take equation (1) and take the log of both sides. Show that $log(y) = log(initial\ amount) + rt$

2) Next let $\beta_0 = log(initialamount)$, let $\beta_1 = r$, and let $x = t$. The model should now look almost just like the log-linear regression model.

3) The only missing piece is u. How would you need to introduce u into equation (1) to get it to be additive in the log-linear model, so it matches that model exactly?

4) What is the interpretation for β_0 in a log-linear model? That is, if we estimated $log(y) = 5 + 0.1x + u$, how would you interpret the 5?

When x = 0, we'd expect y to be ___.

5) What is the interpretation for β_1 in a log-linear model? If we estimated $log(y) = 5 + 0.1x + u$, how would you interpret the 0.1?

When x increases by 1, we'd expect y to increase by ___.

Hint: to solve this question, start with the fact that you know log(y) increases by 0.1, so if y goes from y_1 to y_2, then $log(y_2) - log(y_1) = 0.1$, so $log(\frac{y_2}{y_1}) = 0.1$. Your final answer should look like this: $y_2 - y_1 = ay_1$ for some number a. $a \times 100$ is the percent by which y is expected to increase given a one-unit increase in x.

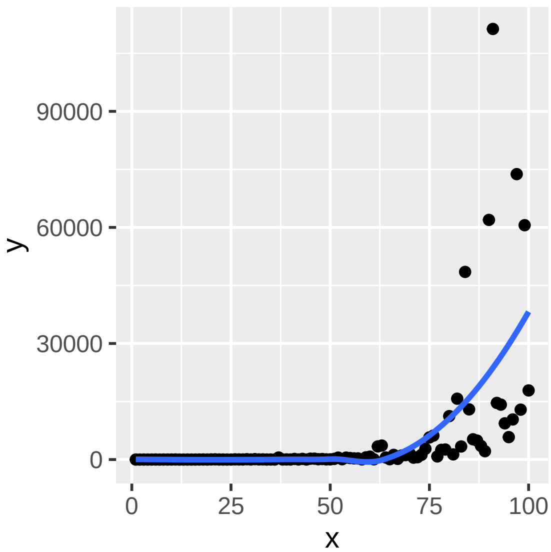
It turns out that there's a simple trick for interpreting β_1: since $r = \beta_1$, the interpretation of β_1 is the same as the interpretation for r: when t increases by 1, y increases by 100 * r %. That's not exactly correct, but it's a close approximation.

Here I'll simulate some data that comes from an exponential growth process, I'll visualize it, I'll visualize it with a log transformation on y, and I'll fit a log-linear model to it.

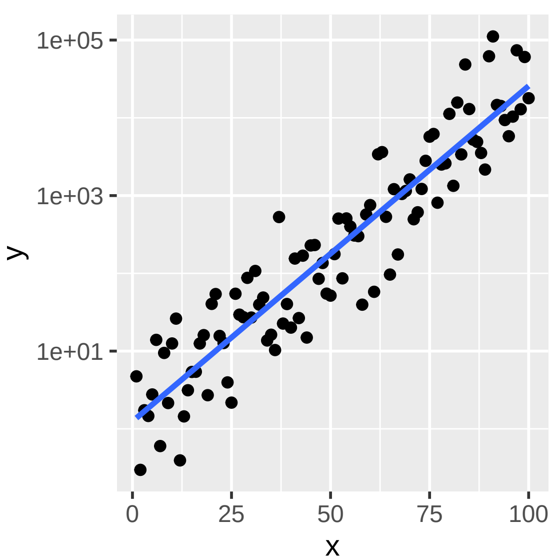
```
library(tidyverse)

exponential_growth_data <- tibble(
```

```
  x = 1:100,
  y = exp(.1 * x + rnorm(n = 100, sd = 1))
  )
```

```
exponential_growth_data %>%
  ggplot(aes(x = x, y = y)) +
  geom_point() +
  geom_smooth(se = FALSE)
```




```
exponential_growth_data %>%
  ggplot(aes(x = x, y = y)) +
  geom_point() +
  geom_smooth(se = FALSE, method = lm) +
  scale_y_log10()
```




```
exponential_growth_data %>%
  lm(log(y) ~ x, data = .) %>%
  broom::tidy()
```

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic  p.value
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)   0.222    0.225       0.984 3.27e- 1
## 2 x             0.0993   0.00387    25.7   2.72e-45
```

Part 2: Log-log

A log-log model has logged terms both on the left and right hand sides: $log(y_i) = \beta_0 + \beta_1 log(x_i) + u_i$.

What kind of process yields this kind of relationship? Consider a constant elasticity demand curve, where the elasticity ε is the percent change in Q_d corresponding to a 1 percent change in price:

$$Q_d = \beta_0 P^{\beta_1} \tag{2}$$

6) Which parameter represents the elasticity ε? Fill in the blanks in the proof below:

$$\begin{aligned}
\varepsilon &= \frac{\%\Delta Q_d}{\%\Delta P} \\
&= \frac{\frac{\partial Q}{Q}}{\frac{\partial P}{P}} \\
&= \frac{\partial Q}{\partial P} \frac{P}{Q} \\
&= \frac{\partial(\beta_0 P^{\beta_1})}{\partial P} \frac{P}{Q} \\
&= (\qquad\qquad\qquad) \frac{P}{Q} \\
&= (\qquad\qquad\qquad) \frac{P}{(\qquad\qquad\quad)} \\
&= \beta_1
\end{aligned}$$

7) Take logs of both sides of equation (2) and change Q to y and P to x to see that we (almost) have the log-log regression model.

8) The only missing piece is u. How would you need to introduce u into equation (2) to get it to be additive in the log-log model, so it matches that model exactly?

β_1 has the same interpretation as an elasticity: it's the expected percent change in y corresponding to a 1 percent change in x. So if we estimate the model and get:

$$log(y) = .25 + .72 log(x)$$

We'd say that a 1 percent increase in x is associated with a .72 percent increase in y. Notice for log-linear, we multiply by 100, but for log-log, we do not.