

STATS 211

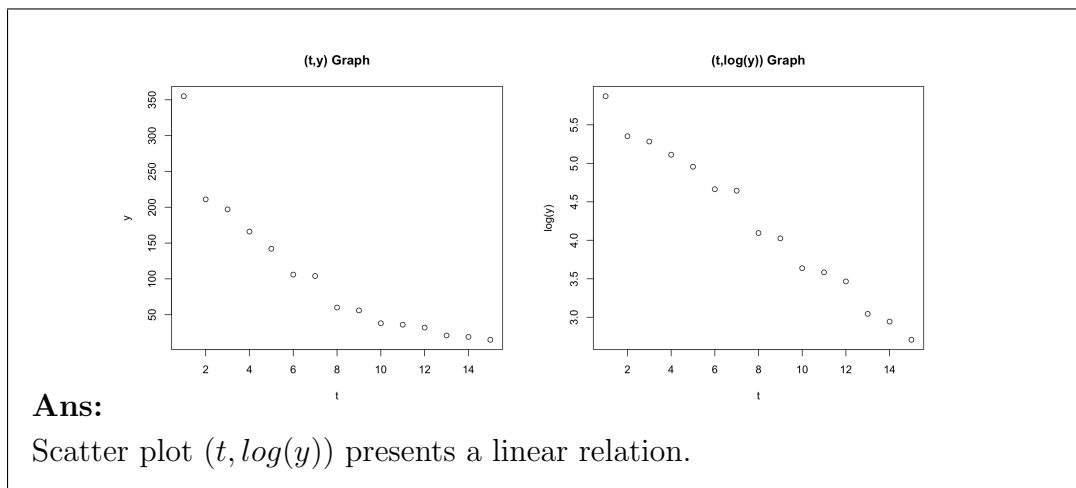
Homework 1

January 30, 2025

Samuel Molero
samueljosemolero@tamu.edu
Section: 501

1. Q1?

- (a) Construct the (t, y) and $(t, \log(y))$ scatter plot. Which scatter plot suggests a linear relationship?



- (b) Construct a predictive equation for the bacteria count Y at time t .

Ans:

```
# R code snippet  
# lm_fit <- lm(y_log~t)  
# summary(lm_fit)
```

Using the command above will provide both the intercept and slope for the linear regression equation using the log form values of Y . Which results in the following equation:

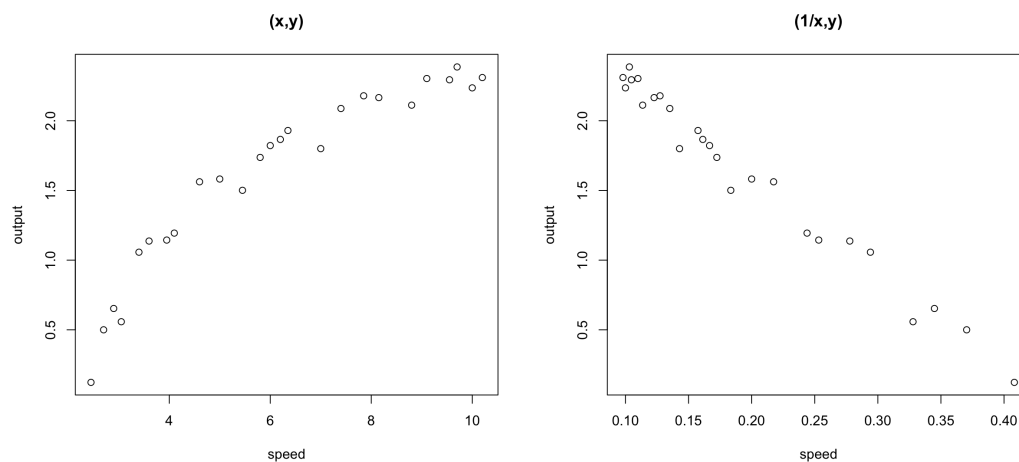
$$\log(\hat{y}) = 5.9732 - 0.2184t$$

Solving for the predictive model for bacteria count Y at time t results in:

$$y = e^{5.9732 - 0.2184t}$$

2. Q2?

Ans:



(a)

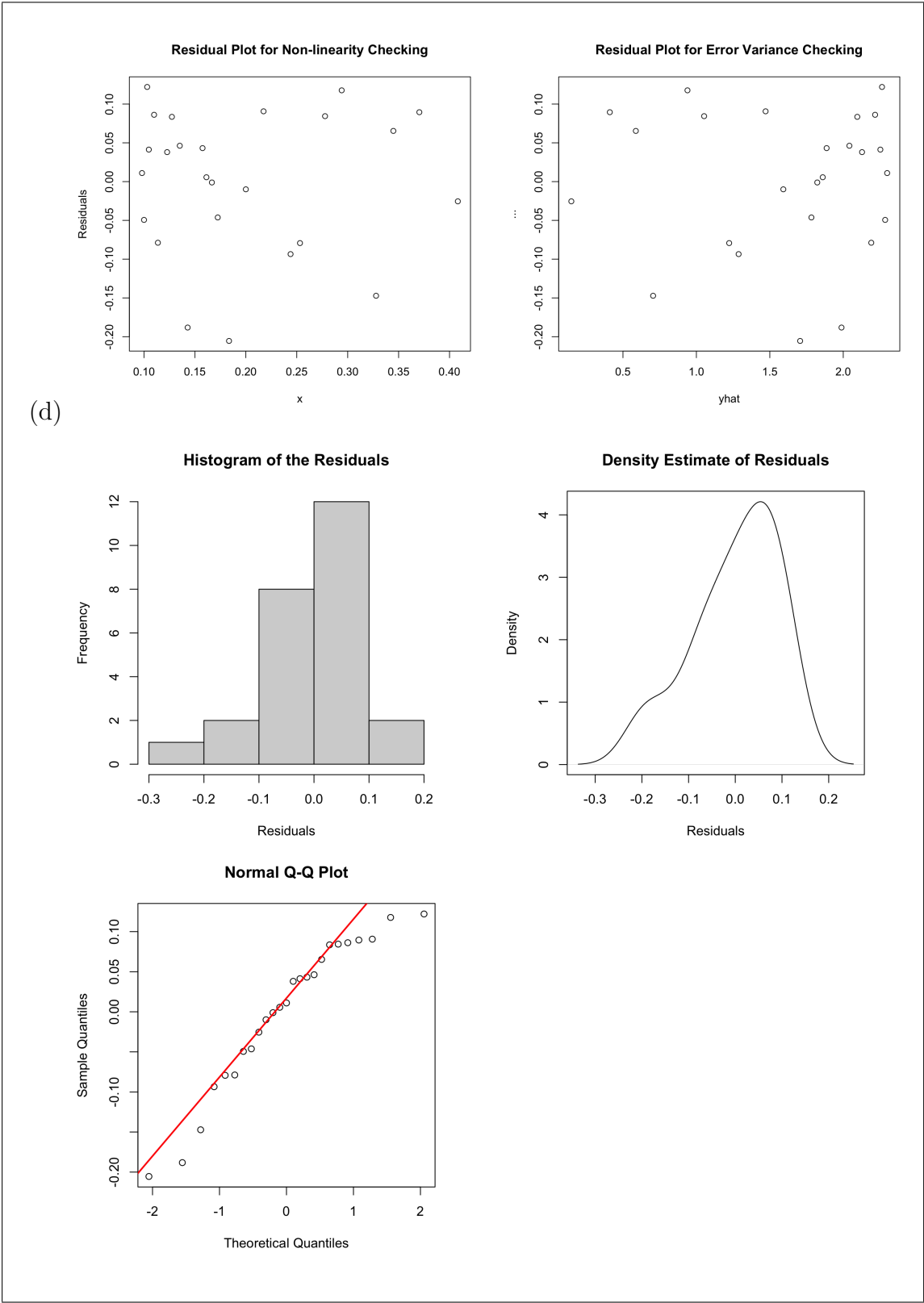
Scatter plot $(1/x, y)$ suggest a linear relation.

(b)

```
#lim_fit <-lm(output~speed2)
#summary(lim_fit)
```

Using the transformed data from the second graph and the commands above, the linear regression is the following: $\hat{y} = 2.9789 - 6.935/x$

(c) Using Wind Speed of 8: $\hat{y} = 2.9789 - 6.935/(8) = 2.1120$



(e) The R^2 obtained equals 0.9800, meaning that 98% of the variation in wind speed's output is explained by linear regression.

(f) `# confint(lim_fit, level=0.99)`

$\beta_1 = [-7.514076 - 6.355019]$, there is 99% confidence that the true slope lies between the given values.

(g) `predict(lm(output~speed2), data.frame(speed=3.2),
interval = "confidence", level=0.95)`

Given that the wind speed is 3.2, there is 95% confidence that the true value lies within the interval [0.74911, 0.87451]

(h) `predict(lim_fit, data.frame(speed=9.05),
interval = "prediction", level=0.95)`

Given that the particular speed at a specific windmill is 9.05, the 95% prediction interval for the output lies within the interval [2.0105, 2.4147]