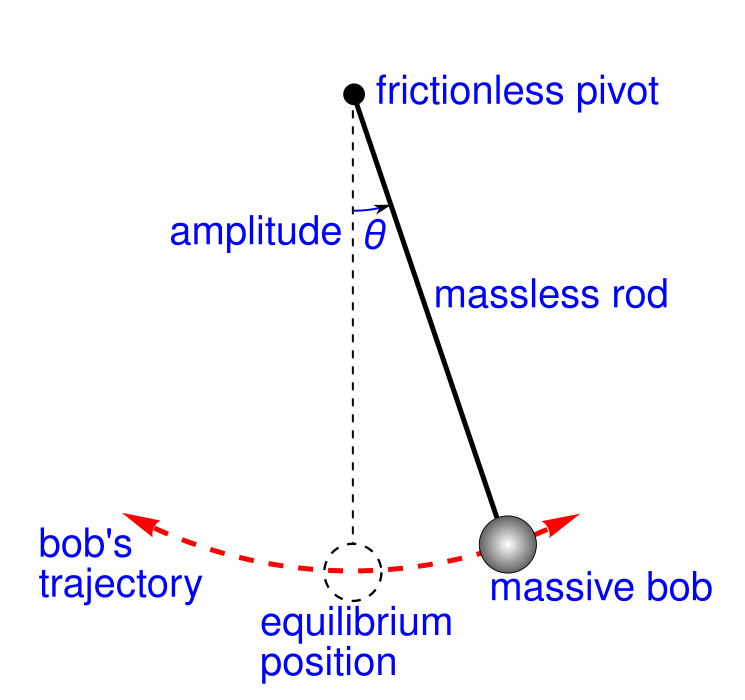
How does a pendulum’s length affect its time period?

Heather Wang, Qihan Liu, Runxi Yu, Vincent Bae

## Introduction

In this experiment, we investigate whether changing the length of a pendulum’s string causes changes in its period (the time it takes to complete one full swing back and forth).



Simple gravity pendulum

Theory suggests that for sufficiently small amounts angles, the pendulum follows where is the period, is the length, and is the acceleration due to gravity. Therefore, the theory suggests that changes in the string’s length do affect the time period.

Therefore, our hypotheses shall be: \* **Null hypothesis:** Changes in the string’s length do not affect the time period. \* **Alternate hypothesis:** Changes in the string’s length do affect the time period.

## Treatment

The treatment in this experiment is the pendulum length, for which we suggested , , , and .

Random assignment is not applicable in this context because the experiment is designed as a deterministic system in a controlled environment; all extraneous variables, such as the mass of the bulb, the gravitational field strength of the earth, were held constant as far as possible. Each length was tested in four trials to reduce measurement errors.

There are some limitations on generalizability. First, although the theoretical formula assumes small-angle oscillations, the initial release angles in this experiment were relatively large. This introduces minor discrepancies due to nonlinear effects at larger angles. Additionally, the results may only generalize to pendulums with similar construction materials and in similar air conditions (e.g., negligible air resistance).

## Subjects

The subjects are the pendulums constructed by tying two 1 CNY coins to a length of string. The same type and number of coins were used throughout the experiment. A piece of string was suspended from a fixed support, and the coins were attached to the end of the string to act as the pendulum bob. The pendulum was released from a fixed angle, and the time it took to complete one period was recorded using a stopwatch.

2 coins were chosen as they are readily available, of a fixed mass, and provide enough mass for the pendulum to swing. The subjects are ideal for the study since the pendulum system is a physical object rather than a social or biological population. They directly represent the physical system being investigated and allow for controlled manipulation of the treatment variable while ensuring other factors are constant. There might be limitations in generalizing the results to all pendulums. The findings will directly apply to pendulums made with similar materials and in similar settings, but may not be generalizable to pendulums with different bobs, materials, or conditions. That said, while the exact timings and results from this experiment is not applicable to all pendulums, the trend and relationship between pendulum length and period are generalizable.

## Analysis

### Data overview

data <- read.csv("data.csv")  
data

## length\_cm time\_10\_periods  
## 1 25 11.01  
## 2 25 10.93  
## 3 25 11.01  
## 4 25 10.94  
## 5 20 10.60  
## 6 20 10.26  
## 7 20 10.16  
## 8 20 10.28  
## 9 15 9.28  
## 10 15 9.18  
## 11 15 9.41  
## 12 15 9.61  
## 13 10 8.40  
## 14 10 8.35  
## 15 10 8.70  
## 16 10 8.76

summary(data)

## length\_cm time\_10\_periods   
## Min. :10.00 Min. : 8.350   
## 1st Qu.:13.75 1st Qu.: 9.075   
## Median :17.50 Median : 9.885   
## Mean :17.50 Mean : 9.805   
## 3rd Qu.:21.25 3rd Qu.:10.682   
## Max. :25.00 Max. :11.010

Treating the length as discrete groups, let us estimate the mean and standard deviation of each group.

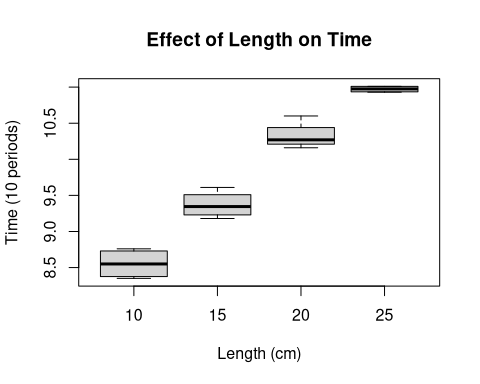
aggregate(time\_10\_periods ~ length\_cm, data = data, function(x) c(mean = mean(x), sd = sd(x), n = length(x)))

## length\_cm time\_10\_periods.mean time\_10\_periods.sd time\_10\_periods.n  
## 1 10 8.55250000 0.20742469 4.00000000  
## 2 15 9.37000000 0.18565200 4.00000000  
## 3 20 10.32500000 0.19070046 4.00000000  
## 4 25 10.97250000 0.04349329 4.00000000

This allows us to see that there is a significant Difference in Means betwen groups, of approximately 0.6 between neighbouring groups, where the standard deviation is roughly around 0.1 to 0.2.

In order to gain a general sense of the data we have obtained, let us plot the data. A box plot is used, to present the information aggregated by group above.

boxplot(time\_10\_periods ~ length\_cm, data = data,  
 main = "Effect of Length on Time",  
 xlab = "Length (cm)",  
 ylab = "Time (10 periods)")



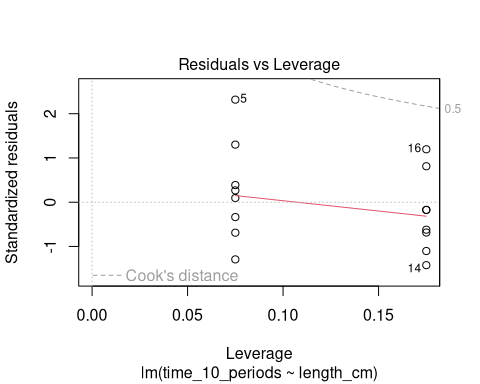
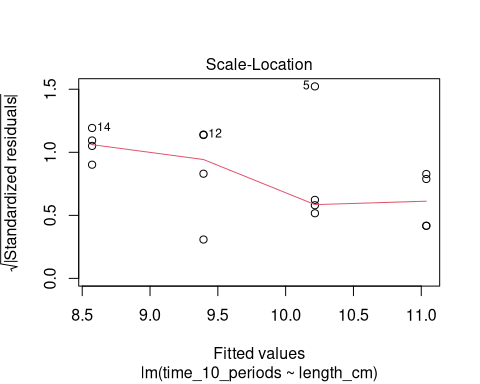
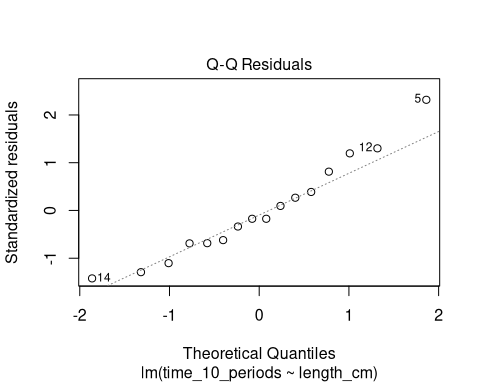
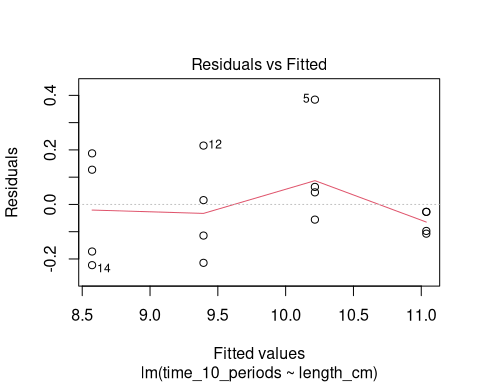
### Linear regression

model <- lm(time\_10\_periods ~ length\_cm, data = data)  
summary(model)

##   
## Call:  
## lm(formula = time\_10\_periods ~ length\_cm, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.22275 -0.10900 -0.02725 0.08000 0.38425   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.929750 0.141582 48.95 < 2e-16 \*\*\*  
## length\_cm 0.164300 0.007707 21.32 4.52e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1723 on 14 degrees of freedom  
## Multiple R-squared: 0.9701, Adjusted R-squared: 0.968   
## F-statistic: 454.5 on 1 and 14 DF, p-value: 4.516e-12

We shall run various diagnostic plots of this linear regression model to confirm whether it is truly a linear regression:

plot(model)



There are insufficient sample points for the diagnostic plots to be accurate, but they seem to be acceptable for this sample size. It is however somewhat unclear what causes the enlarged residues for large theoretical quantities.

Given that a linear regression is an appropriate estimate as suggested by the diagnostic tests (even though the relationship is theoretically supposed to be that of a square root), its -value suggests that it is extremely unlikely to obtain our result (or more extreme results) given that the null hypothesis is true, therefore suggesting with high confidence that the null hypothesis is false and that the alternate hypothesis is true. Therefore: changes in the string’s length do affect the time period.

## T-tests

Another way to confirm the statistical significance of the difference in means is via Welch’s -test, not assuming equal variances.

group\_10 <- data$time\_10\_periods[data$length\_cm == 10]  
group\_15 <- data$time\_10\_periods[data$length\_cm == 15]  
group\_20 <- data$time\_10\_periods[data$length\_cm == 20]  
group\_25 <- data$time\_10\_periods[data$length\_cm == 25]

print(t.test(group\_10, group\_15, var.equal = FALSE))

##   
## Welch Two Sample t-test  
##   
## data: group\_10 and group\_15  
## t = -5.8734, df = 5.9277, p-value = 0.001126  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -1.1590869 -0.4759131  
## sample estimates:  
## mean of x mean of y   
## 8.5525 9.3700

print(t.test(group\_15, group\_20, var.equal = FALSE))

##   
## Welch Two Sample t-test  
##   
## data: group\_15 and group\_20  
## t = -7.1765, df = 5.9957, p-value = 0.000371  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -1.2806739 -0.6293261  
## sample estimates:  
## mean of x mean of y   
## 9.370 10.325

print(t.test(group\_20, group\_25, var.equal = FALSE))

##   
## Welch Two Sample t-test  
##   
## data: group\_20 and group\_25  
## t = -6.6207, df = 3.3113, p-value = 0.005125  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.9428253 -0.3521747  
## sample estimates:  
## mean of x mean of y   
## 10.3250 10.9725

Again, the -values are extremely small, suggesting that it is extremely unlikely to obtain our result (or more extreme results) given that the null hypothesis is true, therefore suggesting with high confidence that the null hypothesis is false and that the alternate hypothesis is true. Therefore: changes in the string’s length do affect the time period.

## Estimated effects

Are effects estimated correctly? Are they described correctly? Is their substantive significance discussed, ideally with a relevant benchmark?

## Statistical Significance

Are p-values presented and discussed correctly? A full score would discuss why the p-value is obtained, i.e. what about the study led to the p-values estimated.

## Informal bibliography

I’m not really sure how to get BibLaTeX working with R Markdown yet, so here’s just an itemized list of references without any particular bibliography format.

* Simple Gravity pendulum, by Chetvorno, public domain, <https://commons.wikimedia.org/w/index.php?curid=5276335>