Marshmallow PVC Pipe Shooters

Jacob Montiel-Bravo, John Linford, and Adam Lenning

Brigham Young University

STAT 230: Analysis of Variance

Professor David Dahl

14 April 2021

Introduction

The principal question of interest is how to make a marshmallow gun more effective, in this case described by the range of the projectiles. This is a topic with vested interest because when you're at war, you need to understand your weapons of choice. When creating a PVC pipe gun, the two major factors are the length and the diameter of the pipe. This is crucial because the mechanics of shooting a marshmallow work through pushing air from the lungs forcefully into a small enough space to exert pressure on the marshmallow being shot. If the space is too small, the force of the breath will not be enough to overcome the friction of the marshmallow in the pipe. If the space is too large, there isn't enough pressure to generate significant speed. Additionally, a longer barrel will generate more velocity, theoretically. So creating a good combination of factors is important to ensure maximum velocity, and therefore, distance.

We expect that the ½ in pipe will create the most force/velocity/distance (our response variable is distance), and that a longer barrel will push the marshmallow farther. As the PVC pipe gets wider, we expect the power and distance to go down, especially at a longer barrel length, because there isn't as much force being exerted directly on the marshmallow.

Design and Data Collection

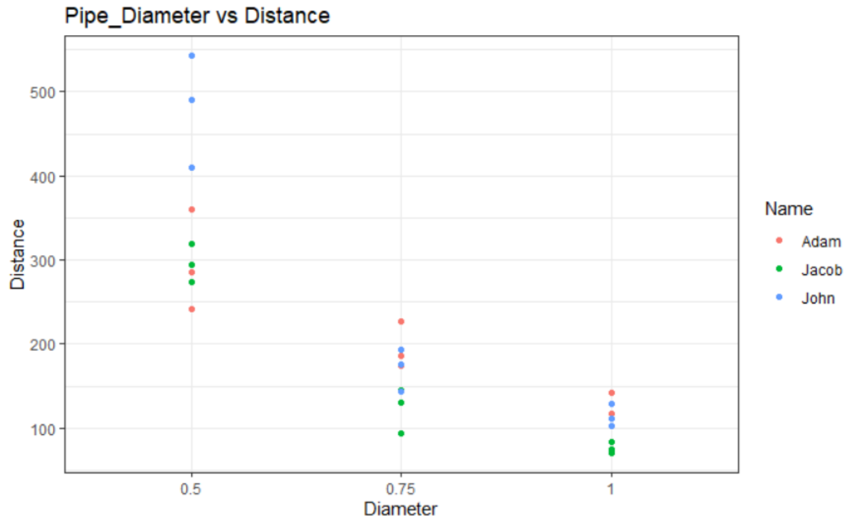
In this experiment, we used a type two complete blocked design (CB[2]). The statistical model is $y_{ijk} = \mu + \beta_i + \alpha_j + \gamma_k + (\alpha\gamma)_{jk} + \varepsilon_{ijk}$, where y_{ijk} is the observation, μ is the overall mean (distance marshmallow traveled), β_i is the effect for the i^{th} block (in this case an individual labeled as "name"), α_j is the effect for the j^{th} level of factor A (the diameter width of the PVC pipe), γ_k is the effect for the k^{th} level of factor C (the length of the PVC pipe), $(\alpha\gamma)_{jk}$ is the interaction between the diameter and the length of the PVC pipes, and ε_{ijk} is the random error effect. The null hypotheses in this experiment are the following: there is no difference in mean distance traveled by marshmallows shot with PVC pipes of different diameters, there is no difference in mean distance traveled by marshmallows shot with PVC pipes of different lengths, and there is no interaction between the diameter of the PVC pipe and the length of the PVC pipe.

Upon conducting a power.anova test, we found that for an n of 27, we had a power level of .52. If we had wanted a power level of .8, we needed 49 individual measurements total, so we roughly should have had double our replicates.

Length/diameter (in)	.5	.75	1
9	1	2	3
12	4	5	6
15	7	8	9

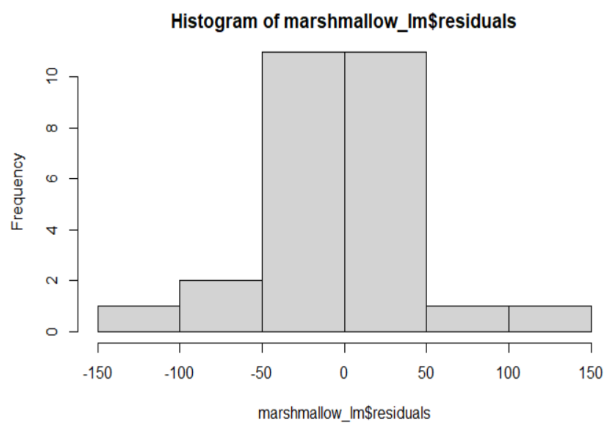
We gathered materials first, with long PVC pipes of three different lengths at .5, .75, and 1 inches respectively. We then measured (with a tape measure) and cut a gun out of each pipe with PVC cutters, to 9, 12, and 15 inches, respectively, in length. We blocked by person, so each individual knelt, and, from a kneeling position, and holding the gun directly in front of them, blew out to launch the marshmallow. The same guns were used for each trial and person. The guns were selected using a random sequence generator (https://www.random.org/sequences/?min=1&max=9&col=1&for mat=html&rnd=new), from 1-9, and corresponded to each gun that had been previously labeled 1-9.

All measurements are in inches. Each person blew each gun three times, then the measurements were averaged out from the three trials of each gun. The surface was a grassy field, we still counted the distance even with weird bounces of the grass, we considered that part of the random chance variation. The marshmallows were placed closest to the mouth side for launching and we used Great Value brand mini marshmallows. Each marshmallow's distance was measured using a tape measure from the location of the shot.

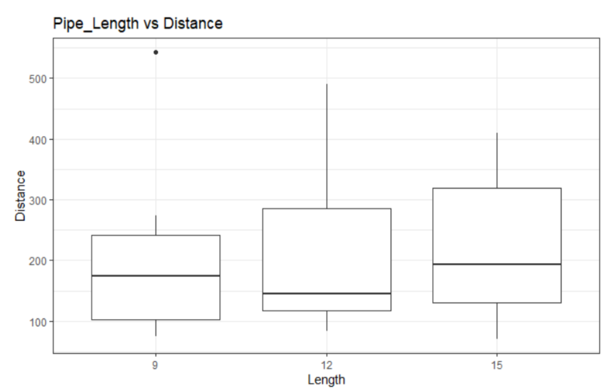


Data Analysis

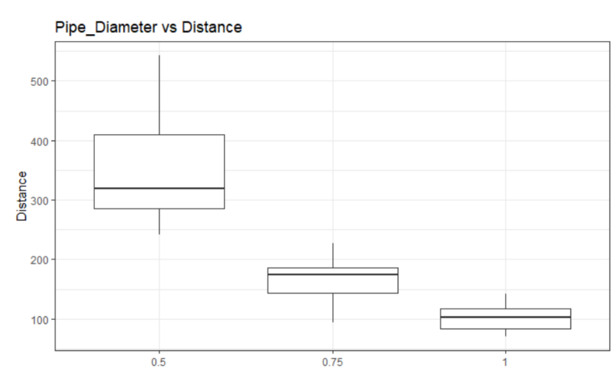
Our assumptions are met, as seen by the histogram of the residuals, we can assume normality and that the error terms will equal zero. We know the results were independent because we randomized the gun selection and usage. The variance seems to be normal and within expected ranges with one potential outlier. We feel confident in moving forward with an ANOVA analysis. Because the residuals are normally distributed and approximately centered at zero, no adjustment is needed.



We found that there is no significance in PVC pipe length or in the interaction between length and diameter. The only factor that was found to be significant was the pipe diameter with a p-value of 5.68e-07 (see the nicely formatted ANOVA table below). From the interaction below, we can see that the .5in pipe trended towards the most distance.



Pipe diameter was found to be the most statistically significant factor in the distance travelled by the marshmallows. We are comparing the three different diameters to discern which PVC pipe diameter creates a significant difference in distance. We found that each comparison that included the .5in PVC pipe was significant.



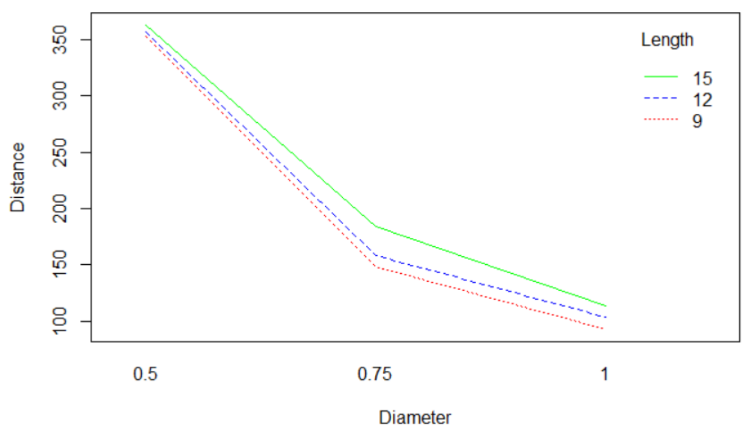
The contrasts, using Tukey's HSD, on the 0.5 in PVC pipe were:

0.5-1: the confidence interval goes from 164.31 to 343.61 and the p-value is .00000029

0.5-0.75: the confidence interval goes from 104.46 to 283.72 and the p-value 0.0000856

Response: Distance (CB2)					
	DF	Sum Sq.	Mean Sq.	F value	P-value
Name	2	36946	18473	4.6914	0.02493
Pipe_Length	2	2258	1129	0.2867	0.75449
Pipe_Diameter	2	317276	158638	40.2873	5.68E-07
Pipe_length:Pipe_Diameter	4	565	141	0.0359	0.99727
Residuals	16	63003	3938		

Interaction Plot:



Conclusion

We found significance only with pipe diameter with a p-value of 5.68e-07. Length and interaction were not significant (p-values 0.75449 and .99727, respectively). These significance results held true when the data were analyzed for a BF[2] and a CB[2]. When the contrast was calculated on the pipe diameter, we found the values measured in the graphic on the right. This is important for future PVC pipe gunmakers because the most crucial factor will always be pipe diameter, and depending on the size of the marshmallows in relation to the pipe size, this will change the potential power of the projectiles. This statistically significant finding does seem to have some causation that can be drawn to other mini marshmallow guns.

```
                diff
0.75-1      59.85184
0.5-1      253.96298
0.5-0.75   194.11113
```

Future studies could compare larger marshmallows with a larger pipe diameter to discern what the proper ratio of marshmallow to gun width will be for maximum distance and power. While there was blocking done for height, it would have been better to have a step stool or a table to standardize the angle of firing. Additionally, another study could work on the effect that air pressure has on the guns and their projectiles, whether a short burst or a drawn-out breath is more effective for maximum efficiency. Other studies could determine marshmallow gun accuracy at different distances and pipe lengths. It could also be useful to have more people/replicates to increase statistical power closer to the .8 range.

Data Sheet

Name	Pipe_Diameter	Pipe_Length	Distance
Jacob	0.5	9	273.6667
Jacob	0.75	9	94
Jacob	1	9	74.3333
Jacob	0.5	12	295
Jacob	0.75	12	145
Jacob	1	12	83.6667
Jacob	0.5	15	318.6667
Jacob	0.75	15	131
Jacob	1	15	70.6667
Adam	0.5	9	241.6667
Adam	0.75	9	175
Adam	1	9	102.3333
Adam	0.5	12	285.6667
Adam	0.75	12	185.6667
Adam	1	12	117.6667
Adam	0.5	15	360.6667
Adam	0.75	15	226.6667
Adam	1	15	142
John	0.5	9	542.6667
John	0.75	9	176
John	1	9	102.6667
John	0.5	12	490.6667
John	0.75	12	144
John	1	12	110.6667
John	0.5	15	409.6667
John	0.75	15	194
John	1	15	128.6667

R-Code

```
---
title: "Final Project"
author: "Jacob Montiel-Bravo, John Linford, Adam Lenning"
date: "4/9/2021"
output: pdf_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)
library(ggplot2)
```

Loading the data

```{r data_load}
marshmallows <- read.csv("Marshmallow-cleaned.csv")
marshmallows$Pipe_Diameter <- as.factor(marshmallows$Pipe_Diameter)
marshmallows$Pipe_Length <- as.factor(marshmallows$Pipe_Length)
marshmallows$Name <- as.factor(marshmallows$Name)
marshmallows
```
Exploratory Data Analysis

```{r exp_analysis}
ggplot(marshmallows, aes(x = Pipe_Diameter, y = Distance, color = Name)) + geom_point() +
theme_bw() + xlab("Diameter") + ylab("Distance") + ggtitle("Pipe_Diameter vs Distance")

ggplot(marshmallows, aes(x = Pipe_Length, y = Distance, color = Name)) + geom_point() +
theme_bw() + xlab("Length") + ylab("Distance") + ggtitle("Pipe_Length vs Distance")

```
Assumptions
```{r}
marshmallow_lm <- lm(Distance ~ Name + Pipe_Length + Pipe_Diameter +
Pipe_Length*Pipe_Diameter, marshmallows)

#Normality
ggplot(marshmallows, aes(x = Pipe_Diameter, y = Distance)) + geom_boxplot() + theme_bw() +
xlab("Diameter") + ylab("Distance") + ggtitle("Pipe_Diameter vs Distance")

ggplot(marshmallows, aes(x = Pipe_Length, y = Distance)) + geom_boxplot() + theme_bw() +
xlab("Length") + ylab("Distance") + ggtitle("Pipe_Length vs Distance")
```
```

````
#Homoscedasticity
hist(marshmallow_lm$residuals)

```

## LM and Anova
```{r}
#BF2
marshmallow_lm <- lm(Distance ~ Pipe_Length + Pipe_Diameter +
Pipe_Length*Pipe_Diameter, marshmallows)
anova(marshmallow_lm)

#CBD2
marshmallow_lm <- lm(Distance ~ Name + Pipe_Length + Pipe_Diameter +
Pipe_Length*Pipe_Diameter, marshmallows)
anova(marshmallow_lm)
```
## Transformation
```{r}

marshmallows$logdistance <- log(marshmallows$Distance)

ggplot(marshmallows, aes(x = Pipe_Diameter, y = logdistance)) + geom_boxplot() +
theme_bw() + xlab("Diameter") + ylab("Distance") + ggtitle("Pipe_Diameter vs Distance")

ggplot(marshmallows, aes(x = Pipe_Length, y = logdistance)) + geom_boxplot() + theme_bw()
+ xlab("Length") + ylab("Distance") + ggtitle("Pipe_Length vs Distance")

marshmallow_lm <- lm(logdistance ~ Name + Pipe_Length + Pipe_Diameter +
Pipe_Length*Pipe_Diameter, marshmallows)
anova(marshmallow_lm)
summary(marshmallow_lm)
```


## Interaction Plot
```{r}
interaction.plot(x.factor = marshmallows$Pipe_Diameter,
                 trace.factor = marshmallows$Pipe_Length,
                 response = marshmallows$Distance,
                 col = c("red", "blue", "green"),
                 trace.label = "Length",
                 ylab = "Distance",
                 xlab = "Diameter")
```
````
````

Power Analysis
```{r}
marshmallow_lm <- lm(Distance ~ Name + Pipe_Length + Pipe_Diameter +
Pipe_Length*Pipe_Diameter, marshmallows)
summary(marshmallow_lm)

#Actual Data
power.anova.test(groups = 3, n = length(marshmallows$Distance), between.var =
var(marshmallows$Distance), within.var = anova(marshmallow_lm)["Pipe_Diameter", "Mean
Sq"], sig.level = 0.05)

#Sample size with power of 0.80
power.anova.test(groups = 3, between.var = var(marshmallows$Distance), within.var =
anova(marshmallow_lm)["Pipe_Diameter", "Mean Sq"], sig.level = 0.05, power = 0.80)

#Power with sample size of 71
power.anova.test(groups = 3, n = 71, between.var = var(marshmallows$Distance), within.var =
anova(marshmallow_lm)["Pipe_Diameter", "Mean Sq"], sig.level = 0.05)

```

Check Contrast
```{r}
#Coefficients produced by R
marshmallow_lm <- lm(Distance ~ Name + Pipe_Length + Pipe_Diameter +
Pipe_Length*Pipe_Diameter, marshmallows)
summary(marshmallow_lm)

#Using Tukey HSD w/ CI and p adjusted
marshmallow.fm <- aov(Distance ~ Pipe_Diameter + Pipe_Length +
Pipe_Diameter*Pipe_Length, data=marshmallows)
TukeyHSD(marshmallow.fm, ordered = TRUE)
```

Means
```{r}
marshmallows %>% group_by(Name, Pipe_Diameter) %>% summarize(mean =
mean(Distance), sd = sd(Distance))
marshmallows %>% group_by(Name, Pipe_Length) %>% summarize(mean = mean(Distance),
sd = sd(Distance))

```