Beer Preferences for Thursday Night Football: A Blanced Incomplete Block Design Experiment

Matthew Murnane   Ethan Newcomb

Abstract

Introduction

American football has immensely grown in popularity in recent years. Along with the growth of a national sport comes an increase in social gatherings and alcoholic beverage enjoyment. Weekly Thursday night football games are broadcast on national television and each Thursday, Matthew's house hosts a watch party along with his roommates. The group enjoys consuming a wide variety of beer, ranging in type and brand. This setting became ideal to test the following research question: What is the beer type that is most favored among the group of friends? Since each individual holds their own inherent biases toward beer, a Balanced Incomplete Block Design (BIBD) is a perfect experimental design. Here, the participants' preferences for beer are blocked, allowing the effects of beer type to be highlighted without the unwanted effect of the nuisance factor. The data collected will be the participants' ratings of beers, which differ in type and flavor. The "Balanced Incomplete" part of the experiment comes from the fact that alcohol impairs judgement, so the amount given to the participants should be limited.

Methods

There are four participants in this study: Zach, Jon, Nolan and Beni. All are male, in their young twenties and beer drinkers. These four participants constituted the entire population of interest in our study.

The constraints of a BIBD design are that when not all treatments can be assigned to each block a subset of those treatments are assigned. To ensure equal precision of estimation of each treatment: Each treatment appears in the same number of block, each treatment has the same replication and there are the same number of pairwise comparisons. A BIBD experiment satisfies the following relation:

$$a \cdot r = b \cdot k$$

$$\text{where, } a = \text{number of treatments in the experiment,}$$

$$r = \text{number of blocks which any one treatment appears,}$$

$$b = \text{number of blocks in the experiment,}$$

$$k = \text{number of treatments per block}$$

We are interested in testing four kinds of beers: A dark, pale ale, IPA, and a generic beer. The brands used for this were such as Guinness for the dark beer, Sierra Nevada for Pale Ale option, Pliny the Elder for IPA and Coors like for the generic beer. That is a $= 4$ treatment options. Our introduction explained the desire to limit the number of beers tasted because alcoholic consumption impairs judgement. To medicate the affect of alcohol on judgement we will limit the number of beers to k $= 3$. With the number of participants being b $= 4$ we will insure that only any given beer appears only r $= 3$ across all blocks. We then have:

$$a \cdot r = b \cdot k$$

$$4 \cdot 3 = 4 \cdot 3$$

$$12 = 12$$

Satisfying the relation proved we can have a valid BIBD experiment.

Everything was randomized before starting the experiment. With simple R code (can be found the the Appendix) and using a `set.seed(530)` we were able to get a correct BIBD set up found in Table 1.

Table 1: Blanced Incomplete Block Design Setup

Participant	I	II	III
1	C	B	D
2	B	C	A
3	D	A	C
4	D	B	A

We then randomized the assignments of beer types to the letters and the order of participants, using the same `set.seed(530)`.The table below is of treatment assignments and order assignments.

Table 2: Randomization of Treatments and Participants

Treatment	Beer	Order	Participant
A	Sierra	1	Zach
B	Coors	2	Jon
C	Guinness	3	Nolan
D	Pliny	4	Beni

The Materials used for the experiment were red solo cups, a pitcher of water, a tablet for recording the ratings, and a blindfold. The red solo cups allowed for the participants to not feel the kind of can the beer came in. This was neccesary to keep them from making guesses that cold affect their judgement.

The experiment was conducted inside of a separate room from the watch party. Participants were blind folded before entry into the room. The blindfold was necessary so that participates did not see what beers were being offered and what beers had already been open. It was best for the experiment that the only sense allowed was the sense of taste. They were led into the room, seated down and told the following:

> You will be offered three beers during the experiment. You will be given a glass of water to cleanse your palate before the tasting begins and another glass of water in between each tasting of beer. After tasting you will rate the beer on a scale from 1 to 10. 1 meaning "I never want to drink this again", 5 meaning "this is an okay beer, and 10 meaning "I want a whole glass of this beer right now". Half points are allowed.

Participants were then given a glass of water to start and then a cup of beer in the kind and order dictated by Tables 1 & 2. After each tasting the participant was given as much time as needed to give the tasted beer an appropriate rating. Ratings were documented on tablet device and transferred into the code found in the Appendix directly after the experiment concluded.

The Design

In summary, we conducted a Balanced Incomplete Block Design where the experimental units are the participants during a specific tasting, the measurement is the rating given by the participant after the tasting, the treatments are the four different kinds of beer and the block are the participants.

A Balanced Incomplete Block Design in analyzed using ANOVA.

The hypothesis:

H_0 : All kinds of beer are like equally, $\mu_i = \mu_{i'} \ \forall i \neq i'$

H_1 : Atleast one beer is not liked equally, $\mu_i \neq \mu_{i'}$ for atleast one pair $i \neq i'$

The model:

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad \text{where } i = 1, 2, 3, 4 \text{ and } j = 1, 2, 3, 4$$
$$\text{and where, } \Sigma_{i=1}^4 \tau_i = 0 \ \text{ and } \ \Sigma_{j=1}^4 \beta_j = 0$$

Note that not all y_{ij} exist.

Results

The ratings for each participant of their three randomly assigned beers are shown in Table 3. Additionally, the mean rating of each beer type and participant are displayed. It can be observed that the beer type with the highest mean rating is Coors (Generic) at 6.17. On the other hand, Guinness (Dark) received the lowest mean rating at 4.33. The range rating between the highest and lowest mean ratings is 1.84, indicating a relatively small difference between mean ratings of beer type. Regarding participant ratings, the highest mean rating resulted from Beni at 6.5, while the lowest mean rating resulted from Zach at 3.67. The range between these two mean ratings is 2.83, indicating a relatively small difference in participant mean ratings.

There appear to be apparent outliers, such as the ratings of 8.5 or 2, which could indicate a significant block effect. However, the ANOVA table results at the end of this section will disprove that notion.

Table 3: Beer Ratings by Participant

Beer	Nolan	Jon	Beni	Zach	Row Means
Sierra(Pale Ale)	NA	5.5	7	2.5	5
Coors(Generic)	6	6	NA	6.5	6.17
Guinness(Dark)	5.5	3.5	4	NA	4.33
Pliny(IPA)	3	NA	8.5	2	4.5
Column Means	4.83	5	6.5	3.67	$\hat{\mu} = 5$

Figure 1 displays a visual representation of Ratings Per Beer. It is relatively difficult to draw a conclusion from this data, given the wide range of data points and their overlapping nature.
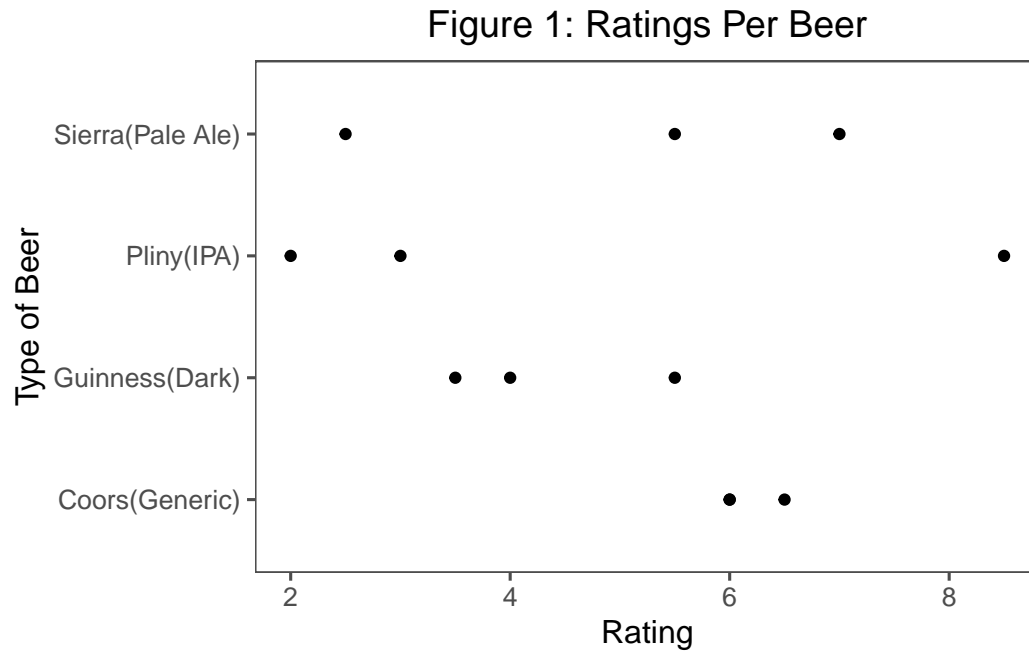
Figure 1: Ratings Per Beer

Figure 2 displays a visual representation of Ratings Per Subject. One possible trend observed in this chart is the different rating tendencies of the participants. Although the data overlaps, a potential trend could be that Beni naturally rates beers higher while Zach naturally rates them lower.
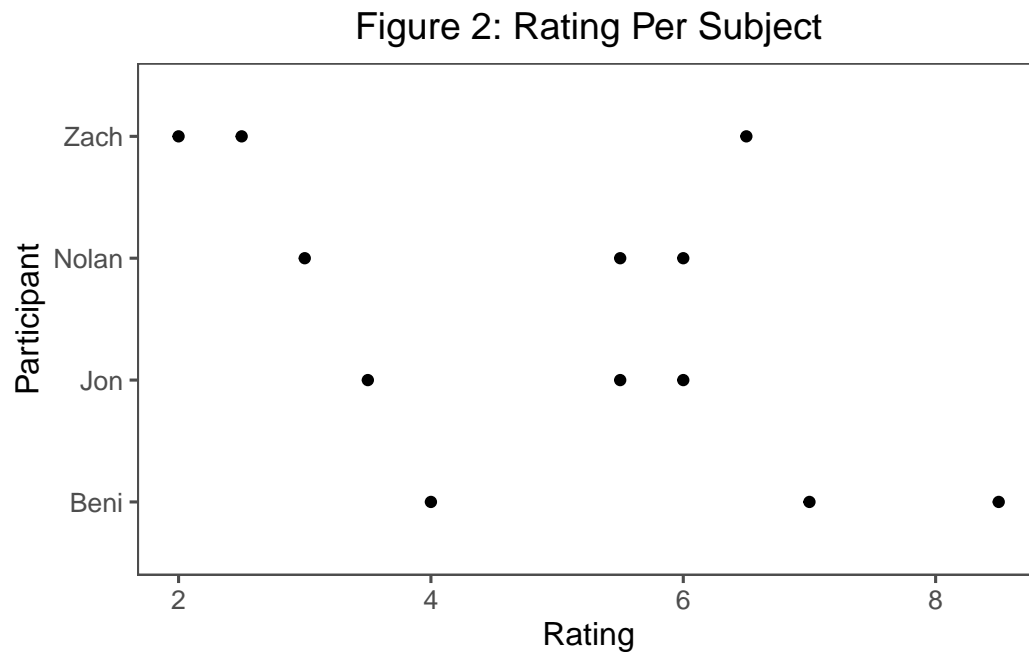
Figure 2: Rating Per Subject

Table 4 displays the ANOVA results from the BIBD design. The p-value for the treatment (type of beer) is 0.40. and the p-value for the block (participant) is 0.34. Meaning that our factor of interest and

blocking factor are both insignificant.

Table 4: ANOVA Table for Linear Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Participant	3	12.17	4.06	1.20	0.40
Beer	3	14.40	4.80	1.42	0.34
Residuals	5	16.94	3.39	NA	NA

Conclusion

Appendix

Code Used

Libraries Used

```
library(tidyverse)
library(ggthemes)
library(tidyr)
library(knitr)
```

Data Code

```
# Data input
beers <- c("Sierra(Pale_ale)", "Coors(Generic)", "Guinness(Dark)", "Pliny(IPA)")
Nolan <- c(NA, 6, 5.5, 3)
Jon <- c(5.5, 6, 3.5, NA)
Benni <- c(7, NA, 4, 8.5)
Zach <- c(2.5, 6.5, NA, 2)

raw_data <- data.frame(beers, Nolan, Jon, Benni, Zach)

# Data Cleaning

pivoted_raw_data <- pivot_longer(raw_data,
                       cols=-beers,
                       names_to = "names",
                       values_to = "rating")

cleaned_data <- pivoted_raw_data %>%
  drop_na(rating)


#Table
col_means <- round(colMeans(raw_data[, -1], na.rm = TRUE), 2)

raw_data_with_col_means <- rbind(raw_data, c("Mean", col_means))

row_means <- round(apply(raw_data[, -1], 1, mean, na.rm = TRUE), 2)

raw_data_with_means <- cbind(raw_data_with_col_means, Row_Mean = c(row_means, NA))
```

```r
raw_data_with_means %>%
  kable(
    caption = "Beer Ratings by Participant",
    col.names = c("Beers", "Nolan", "Jon", "Beni", "Zach", "Row Means"),
    align = "c"
  )
```

Randomization Code

```r
# Computing the BIBD matrix
trts <- c("A", "B", "C", "D")
set.seed(530)
t(replicate(4, sample(trts, 3, replace = FALSE)))

# Randomly matching treatments to the letters
brands <- c("Pliny", "Coors", "Siera", "Guinnes")
shuffled_brands <- sample(brands)
assignments <- data.frame(trts, shuffled_brands)

#Randomizing the order of participants/blocks
boys <- c("Zach", "Jon", "Nolan", "Benni")
rank <- 1:4
shuffled_names <- sample(boys)
order <- data.frame(rank, shuffled_names)
```

Plots

```r
cleaned_data %>%
  ggplot(aes(x = rating,
             y = beers))+
  geom_point(size = 1.5)+
  theme_few()+
  ggtitle("Ratings Per Beer")+
  ylab("Type of Beer")+
  xlab("Rating")

cleaned_data %>%
  ggplot(aes(x = rating,
             y = names))+
  geom_point(size = 1.5)+
```

```
  theme_few()+
  ggtitle("Rating Per Subject")+
  xlab("Subjects")+
  ylab("Rating")
```

Anova Table

```
linear_model <- lm(Rating~Participant+Beer, data = cleaned_data)
anova_table <- anova(linear_model)
rounded_anova_table <- anova_table
rounded_anova_table[] <- lapply(anova_table, function(x) {
  if (is.numeric(x)) round(x, 2) else x
})
kable(rounded_anova_table,
      caption = "ANOVA Table for Linear Model")
```