Analysis of Poker Experiment

Submitted to Dr. Ginger Holt

Haritha Peyyeti MATH 261 B Project May 21st 2015

1. Introduction

Poker is a game of chance in most of the countries, while some think that it should be treated as a game of skill because the final outcome of winning or losing the game primarily depends on the individual who is playing. Although the available studies indicate that skill plays an important role, there is no information regarding the relative importance of chance and skill. In the present project I analyzed the extent to which the influence of poker playing skill is more important than card distribution along with other effect of bet limit on winnings.

2. Experimental Design

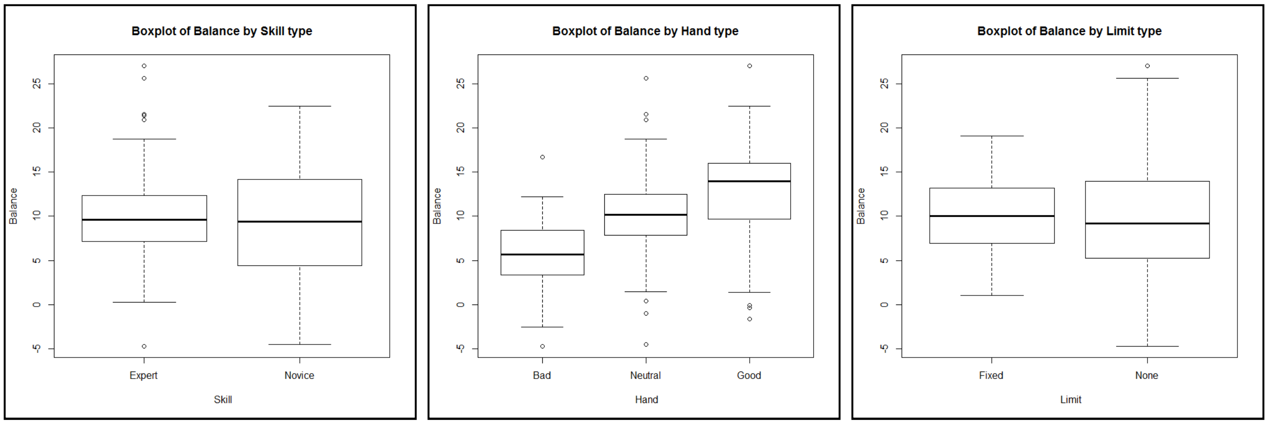
The purpose of this experiment is to determine how the skill of an individual, the card distribution or hand, and the limitation on the bet amount affect the winnings in a poker game. The response variable is final cash balance (in euros). The 3 factors and their levels[1] are:

Factors	Levels		
Skill	Expert	Average	
Hand	Bad	Neutral	Good
Limit	Fixed	None	

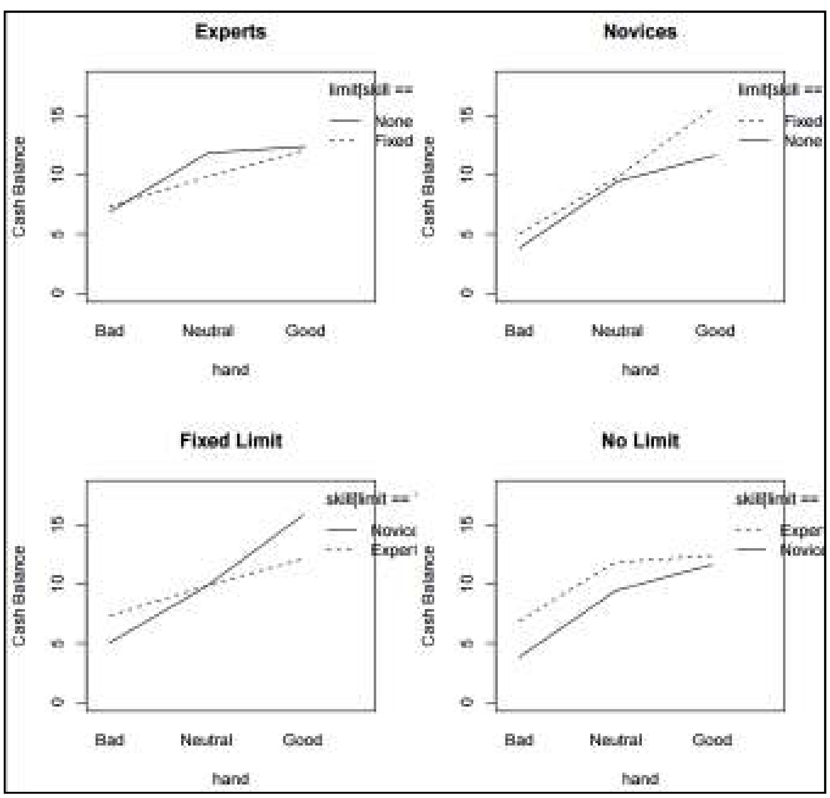
25 different individuals played in each of the 12 conditions (each individual in only one treatment). In total, 150 individuals participated in a "fixed-limit" game variant, and 150 individuals participated in a "no-limit" game variant. Each individual played computer-based hands of the poker variant "Texas Hold'em" for money. In each hand, one of the novice / average players and one expert players received (a) better-than-average cards i.e good box, (b) average cards i.e. neutral box and (c) worse-than-average cards i.e. loser's box. The randomization of the card distribution controlled the factor of chance to determine differences in performance between the average and expert groups. It is to be noted here that this is an experimental design with 25 repeated measurements (i.e. 25 individuals played for each of the 12 treatment combination) and not 25 replicates since no individual plays for more than once. Randomization, one of the basic principles of experimental design is followed. Each novice or expert player was randomly assigned a good, bad or neutral hand. This experiment is a classic example of a 3 factor ANOVA model with mixed levels. Since the levels or treatments were specifically chosen, it is a Fixed Effects model. The conclusions that we draw from this experiment will apply only to the factor levels considered here and cannot be extended to other treatments that are not considered in the analysis.

[1] The Design matrix can be found in the Appendix.

3. Exploratory Data Analysis



Preliminary analysis shows that the average cash balance left does not differ much within the two treatment groups ('Expert' and 'Novice') of skill type and two treatment groups of bet limit ('Fixed' and 'None'). However, cash balances are clearly higher for a 'Good' hand type than for a 'Neutral' than for a 'Bad' hand type. This plot suggests that hand type or the card distribution is a major factor in deciding the winnings. We need to understand the interaction between the 3 factors and the corresponding levels.



Fixed Bet Limit	Skill		
Hand	Novice	Expert	Overall
Bad	€ 5.05	€ 7.33	€ 6.19
Neutral	€ 9.80	€ 9.85	€ 9.83
Good	€ 15.80	€ 12.12	€ 13.96
Overall	€ 10.22	€ 9.77	€ 9.99

No Bet Limit	Skill		
Hand	Novice	Expert	Overall
Bad	€ 3.84	€ 6.89	€ 5.36
Neutral	€ 9.45	€ 11.85	€ 10.65
Good	€ 11.63	€ 12.39	€ 12.01
Overall	€ 8.31	€ 10.38	€ 9.34

From the above interaction plots and tables, we notice the interaction effect between 'hand' type and 'skill' type is significant, while the interaction between 'hand' type and bet 'limit' is not quite significant. Also, we notice that for a bad or loser's card distribution, an expert ends up with a higher cash balance than a novice. However, it is little surprising that a novice player is left with higher average cash balance than an expert player under a good or winner's card distribution type. The exploratory data analysis of main effects and interaction effects signals that the 'hand' type may be the significant main effect and the other main effects may not be significant. Also, the interaction effect between 'hand' and 'skill' might be significant, while the other interaction effects have to be further analyzed.

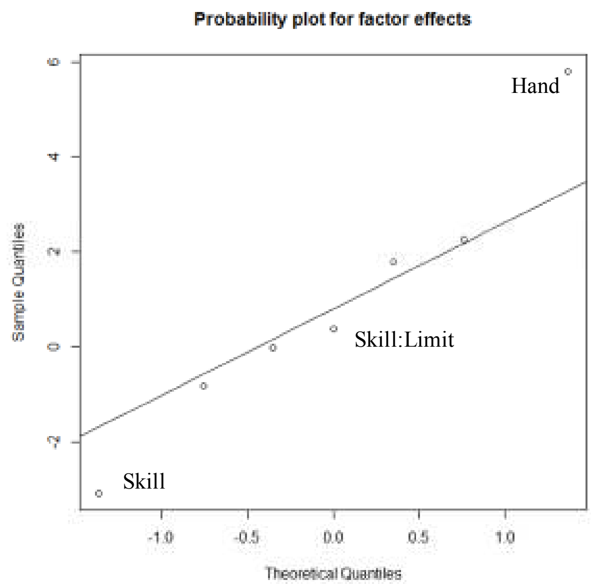
4. ANOVA Model

The different games by each of the 25 individuals in a treatment group are not replicates, but repeated observations by different individuals on the same experimental levels. It is not a good idea to use the data from the experiment as if there were 25 replicates of a 3 factor mixed level design because the estimate of error may not reflect the treatment variation. Therefore, we analyze average and standard deviation of the cash balance left.

1. Average of Cash Balance: Model Fitting

```
> poker.avg = aggregate(cash ~ skill*hand*limit, poker, mean)
> fit.avg <- lm(cash~-1 + skill*hand*limit, data=poker.avg)
> anova(fit.avg)
Analysis of Variance Table

Response: cash
                 Df Sum Sq Mean Sq  F value    Pr(>F)    
skill             1 981.17  981.17 625.2923 1.908e-06 ***
hand              1 235.86  235.86 150.3133 6.388e-05 ***
limit             1   1.02    1.02   0.6475   0.45754    
skill:hand        1   1.04    1.04   0.6656   0.45168    
skill:limit       1  20.59   20.59  13.1239   0.01518 *  
hand:limit        1   1.95    1.95   1.2439   0.31543    
skill:hand:limit  1   0.00    0.00   0.0006   0.98097    
Residuals         5   7.85    1.57                       
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the model on average cash balance, it is observed that 'skill', 'hand' and 'skill':'limit" interaction effects are significant.



Refitting the reduced model with only the significant effects:

```
> fitavg.red1 <- lm(cash~hand+ skill*limit,data=poker.avg)
> anova(fitavg.red1)
Analysis of Variance Table

Response: cash
            Df  Sum Sq Mean Sq F value    Pr(>F)
hand         1 103.908 103.908 44.6163 0.0002827 ***
skill        1   1.967   1.967  0.8445 0.3886914
limit        1   1.267   1.267  0.5441 0.4847123
skill:limit  1   4.764   4.764  2.0457 0.1957165
Residuals    7  16.302   2.329
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> summary(fitavg.red1)

Call:
lm(formula = cash ~ hand + skill * limit, data = poker.avg)

Residuals:
     Min       1Q   Median       3Q      Max
-1.58915 -0.95996 -0.09831  1.14935  1.97898

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.0230     4.5356  -0.226 0.828003
hand          3.6040     0.5396   6.680 0.000283 ***
skill         2.9709     2.7862   1.066 0.321681
limit         3.1307     2.7862   1.124 0.298225
skill:limit  -2.5204     1.7622  -1.430 0.195716
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.526 on 7 degrees of freedom
Multiple R-squared:  0.8728, Adjusted R-squared:  0.8002
F-statistic: 12.01 on 4 and 7 DF,  p-value: 0.002973
```

We notice that only the main effect "Hand" is significant. Refitting the model, with factors 'hand', 'skill' and interaction between the two, we see a better model fit (R-squared adj = 0.85)

```
> fitavg.red2 <- lm(cash~skill*hand,data=poker.avg)
> anova(fitavg.red2)
Analysis of Variance Table

Response: cash
           Df  Sum Sq Mean Sq F value    Pr(>F)    
skill       1   1.967   1.967  1.1375   0.31730    
hand        1 103.908 103.908 60.1009 5.472e-05 ***
skill:hand  1   8.503   8.503  4.9181   0.05739 .  
Residuals   8  13.831   1.729                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

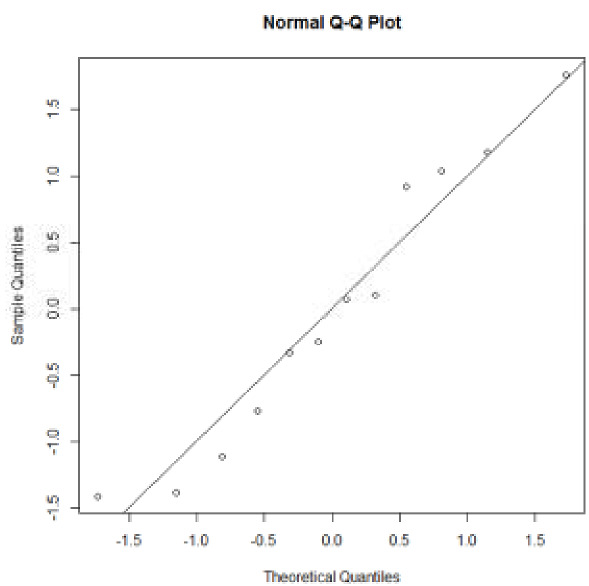
```
> summary(fitavg.red2)

Call:
lm(formula = cash ~ skill * hand, data = poker.avg)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.2665 -0.5461 -0.1952  0.4524  1.9031 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)   9.8587     3.1757   3.104   0.0146 *
skill        -4.9335     2.0085  -2.456   0.0395 *
hand          0.5111     1.4701   0.348   0.7371  
skill:hand    2.0619     0.9298   2.218   0.0574 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.315 on 8 degrees of freedom
Multiple R-squared:  0.8921,	Adjusted R-squared:  0.8517 
F-statistic: 22.05 on 3 and 8 DF,  p-value: 0.0003186
```

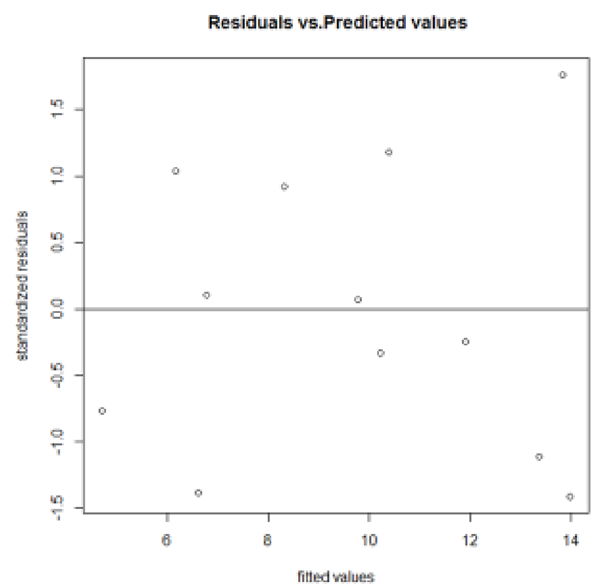
Refitting the model, with only 'hand' main effect, we noticed adj R-squareed of 0.79 and significant p-value indicating that the model is adequate.

```
> fitavg.red3 <- lm(cash~hand,data=poker.avg)
> anova(fitavg.red3)
Analysis of Variance Table

Response: cash
          Df  Sum Sq Mean Sq F value    Pr(>F)    
hand       1 103.908  103.91  42.759 6.563e-05 ***
Residuals 10  24.301    2.43                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> summary(fitavg.red3)

Call:
lm(formula = cash ~ hand, data = poker.avg)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.22288 -1.04686 -0.04143  0.93672  2.52922 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.4585     1.1906   2.065   0.0658 .  
hand          3.6040     0.5511   6.539 6.56e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.559 on 10 degrees of freedom
Multiple R-squared:  0.8105, Adjusted R-squared:  0.7915 
F-statistic: 42.76 on 1 and 10 DF,  p-value: 6.563e-05
```

2. Average of Cash Balance : Model Diagnostics



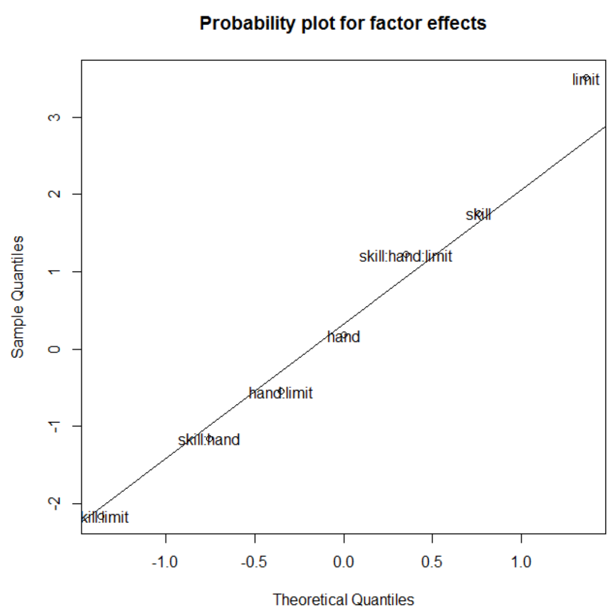


The fit2 model (with ‘hand’, ‘skill’ and ‘hand’: ‘skill’ interaction) gives a Lack-of-fit p-value of 0.31, and an adjusted R^2= 0.85 which suggests a good fit. In addition to this, residual plots appear to be satisfactory. The estimated MSE is found to be 1.73. There appears to be no violations to normality or variance assumptions.

3. **Standard deviation of Cash Balance : Model fitting**

```
> poker.sd = aggregate(cash ~ skill*hand*limit, poker, sd)
> fit.sd <- lm(cash~-1 + skill*hand*limit, data=poker.sd)
> anova(fit.sd)
Analysis of Variance Table

Response: cash
                  Df  Sum Sq Mean Sq  F value    Pr(>F)    
skill              1 156.905 156.905 748.2629 1.222e-06 ***
hand               1  12.561  12.561  59.9008 0.0005755 ***
limit              1  31.628  31.628 150.8292 6.335e-05 ***
skill:hand         1   1.368   1.368   6.5233 0.0510123 .  
skill:limit        1   0.008   0.008   0.0381 0.8528457    
hand:limit         1   1.835   1.835   8.7524 0.0315758 *  
skill:hand:limit   1   2.458   2.458  11.7226 0.0187603 *  
Residuals          5   1.048   0.210                       
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

From the model on standard deviation of cash balances left, we notice that all the three main effects- skill, hand, limit and 3 interaction effects - skill:hand, skill:hand:limit, hand:limit are significant.

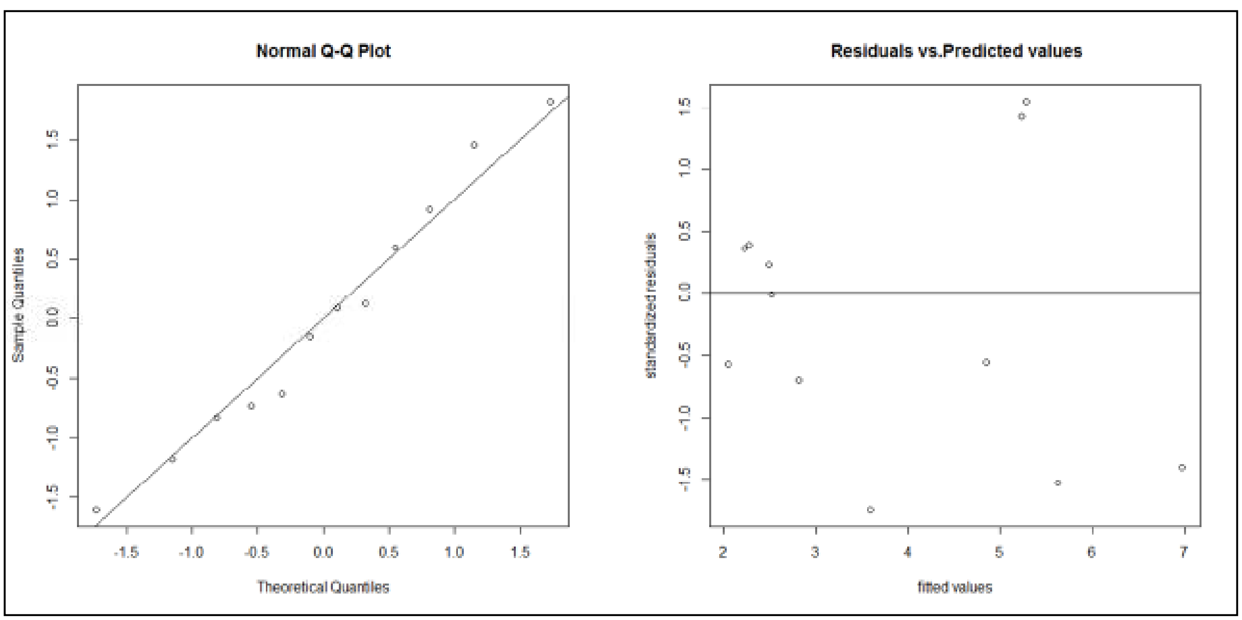


Refitting the model with only significant effects and dropping the rest, we finally selected a model with limit and hand:limit effects.

```
> fitred2.sd <- lm(cash~limit+hand*limit, data=poker.sd)
> anova(fitred2.sd)
Analysis of Variance Table

Response: cash
           Df  Sum Sq Mean Sq F value    Pr(>F)
limit       1 24.8872 24.8872 65.1142 4.104e-05 ***
hand        1  1.2889  1.2889  3.3723   0.10362
limit:hand  1  3.0135  3.0135  7.8844   0.02292 *
Residuals   8  3.0577  0.3822
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> summary(fitred2.sd)

Call:
lm(formula = cash ~ limit + hand * limit, data = poker.sd)

Residuals:
     Min       1Q   Median       3Q      Max
-1.01015 -0.09388  0.08502  0.42949  0.59543

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.3843     1.4932   1.597   0.1490
limit         0.4252     0.9444   0.450   0.6644
hand         -1.4398     0.6912  -2.083   0.0708 .
limit:hand    1.2275     0.4372   2.808   0.0229 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6182 on 8 degrees of freedom
Multiple R-squared:  0.9052, Adjusted R-squared:  0.8696
F-statistic: 25.46 on 3 and 8 DF,  p-value: 0.0001912
```

4. Standard deviation of Cash Balance : Model fitting



The reduced model (with limit and hand: limit interaction) gives a Lack-of-fit p-value of 0.53, and an adjusted R^2= 0.86 which suggests a good fit. In addition to this, residual plots appear to be satisfactory. The estimated MSE is found to be 0.382. There appears to be no gross violations to normality or variance assumptions.

5. Conclusion

ANOVA results showed that expert players did not outperform average or novice players in terms of final cash balance. Instead, card distribution was the significant factor for successful poker playing. However, expert players were better able to minimize losses and end up with higher cash balances when confronted with disadvantageous worse-than-average cards. No significant differences were observed between the game variants. Furthermore, standard deviation analysis confirms that bet limit affects the standard deviation in the cash balances. To summarize, the findings of the poker experiment indicate that poker should be considered more as a game of chance than as a game of skill, at least under certain basic conditions.

6. References

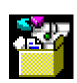
Montgomery D C 2013 Design and Analysis of Experiments (8th edn). John Wily and Sons Inc, Singapore.

G. Meyer, M. von Meduna, T. Brosowski, T. Hayer (2012). "Is poker a Game of Skill or Chance? A Quasi-Experimental Study," Journal of Gambling Studies, Online First DOI 10.1007/s10899-012-9327-8

Appendix A: Design Matrix Structure

> poker

	skill	hand	limit	cash
1	1	1	1	4.00
2	1	1	1	5.55
3	1	1	1	9.45
4	1	1	1	7.19
5	1	1	1	5.71
6	1	1	1	5.32
7	1	1	1	8.52
8	1	1	1	4.06
9	1	1	1	4.97
10	1	1	1	6.18
11	1	1	1	6.55
12	1	1	1	9.73
13	1	1	1	10.80
14	1	1	1	2.07
15	1	1	1	8.69
16	1	1	1	8.47
17	1	1	1	7.00
18	1	1	1	10.67
19	1	1	1	11.04
20	1	1	1	10.41
21	1	1	1	10.05
22	1	1	1	2.92
23	1	1	1	8.39
24	1	1	1	8.36
25	1	1	1	7.13
26	1	1	2	5.52
27	1	1	2	8.28
28	1	1	2	0.27
29	1	1	2	11.49
30	1	1	2	-4.68
31	1	1	2	12.22
32	1	1	2	6.24
33	1	1	2	5.10
34	1	1	2	3.76
35	1	1	2	1.00
36	1	1	2	9.52


poker_exp.txt

Complete design matrix can be found here:

Appendix B: R Code

```
setwd("C:/bujji/SEM2/MATH 261B/Data")

poker <- read.table("poker_num.txt",header=F,col.names=c("skill","hand","limit","cash"))

attach(poker)

skill <- factor(skill); levels(skill) <- c("Expert","Novice")

hand <- factor(hand); levels(hand) <- c("Bad","Neutral","Good")

limit <- factor(limit); levels(limit) <- c("Fixed","None")

mean(cash)

(tapply(cash,list(skill,hand,limit),mean)); (tapply(cash,list(skill,hand),mean))

(tapply(cash,list(skill,limit),mean)); (tapply(cash,list(hand,limit),mean))

(tapply(cash,skill,mean)); (tapply(cash,hand,mean)); (tapply(cash,limit,mean))

plot(cash~skill, main = "Boxplot of Balance by Skill type",xlab = " Skill", ylab= "Balance")

plot(cash~hand, main = "Boxplot of Balance by Hand type",xlab = " Hand", ylab= "Balance")

plot(cash~limit, main = "Boxplot of Balance by Limit type",xlab = " Limit", ylab= "Balance")

par(mfrow=c(2,2))

interaction.plot(hand[skill=="Expert"],limit[skill=="Expert"],cash[skill=="Expert"],
        ylim=c(0,18),xlab="Hand", ylab="Cash Balance",main="Experts",trace.limit="lim")

interaction.plot(hand[skill=="Novice"],limit[skill=="Novice"],cash[skill=="Novice"],
        ylim=c(0,18),xlab="hand", ylab="Cash Balance",main="Novices")

interaction.plot(hand[limit=="Fixed"],skill[limit=="Fixed"],cash[limit=="Fixed"],
        ylim=c(0,18),xlab="hand", ylab="Cash Balance",main="Fixed Limit")

interaction.plot(hand[limit=="None"],skill[limit=="None"],cash[limit=="None"],
        ylim=c(0,18),xlab="hand", ylab="Cash Balance",main="No Limit")


poker.aov1 <- aov(cash ~ as.factor(skill)*as.factor(hand)*as.factor(limit), data=poker.avg)

summary(poker.aov1)

################Avg model##################

poker.avg = aggregate(cash ~ skill*hand*limit, poker, mean)

fit.avg <- lm(cash~-1 + skill*hand*limit, data=poker.avg)

anova(fit.avg)

coef(fit.avg)
```

```
plotavg = qqnorm(coef(fit.avg), main="Probability plot for factor effects")

qqline(coef(fit.avg))

text(plotavg$x, plotavg$y, names(plotavg$y))

fitavg.red1 <- lm(cash~hand+ skill*limit,data=poker.avg)

anova(fitavg.red1)

summary(fitavg.red1)

fitavg.red2 <- lm(cash~skill*hand,data=poker.avg)

anova(fitavg.red2)

summary(fitavg.red2)

fitavg.red3 <- lm(cash~hand,data=poker.avg)

anova(fitavg.red3)

summary(fitavg.red3)

qqnorm(rstandard(fitavg.red1))

abline(0,1)

plot(fitted(fitavg.red1),rstandard(fitavg.red1), main = " Residuals vs.Predicted values", xlab="fitted values",
ylab="standardized residuals")

abline(0,0)

qqnorm(rstandard(fitavg.red2))

abline(0,1)

plot(fitted(fitavg.red2),rstandard(fitavg.red2), main = " Residuals vs.Predicted values", xlab="fitted values",
ylab="standardized residuals")

abline(0,0)

qqnorm(rstandard(fitavg.red3))

abline(0,1)

plot(fitted(fitavg.red3),rstandard(fitavg.red3), main = " Residuals vs.Predicted values", xlab="fitted values",
ylab="standardized residuals")

abline(0,0)


####sd##################

poker.sd = aggregate(cash ~ skill*hand*limit, poker, sd)

fit.sd <- lm(cash~-1 + skill*hand*limit, data=poker.sd)

anova(fit.sd)

coef(fit.sd)
```

```
summary(fit.sd)

plotsd = qqnorm(coef(fit.sd), main="Probability plot for factor effects")

qqline(coef(fit.sd))

text(plotsd$x, plotsd$y, names(plotsd$y))


fitred1.sd <- lm(cash~skill+hand+limit+skill*hand+skill*hand*limit+hand*limit, data=poker.sd)

anova(fitred1.sd)

summary(fitred1.sd)

coef(fitred1.sd)

fitred2.sd <- lm(cash~limit+hand*limit, data=poker.sd)

anova(fitred2.sd)

summary(fitred2.sd)

coef(fitred2.sd)

qqnorm(rstandard(fit.sd))

abline(0,1)

plot(fitted(fit.sd),rstandard(fit.sd), main = " Residuals vs.Predicted values", xlab="fitted values", ylab="standardized residuals")

abline(0,0)

qqnorm(rstandard(fitred2.sd))

abline(0,1)

plot(fitted(fitred2.sd),rstandard(fitred2.sd), main = " Residuals vs.Predicted values", xlab="fitted values", ylab="standardized residuals")

abline(0,0)

dev.off()
```