# Free Throw Shooting Routine Factor Screening

Alex Salce, STAT571B

Spring 2023

# 1 Introduction

As someone who enjoys playing basketball recreationally, constructing an experimental design that may give objective insight to factor(s) that could potentially help me improve an aspect of my play was of interest. Being that I would be the only subject, it made sense to design an experiment based on the free throw. The free throw is an individual play with discrete outcomes (make or miss), and prior to a free throw shot being taken, some individual "routine" is executed; a combination of form, time, and performative actions that are widely percieved to affect how well one shoots. Even to the highest levels of competitive basketball, many believe the routine to have an impact upon the ability to shoot free throws in some capacity. Some famous examples of free throw routines include Giannis Antetokounmpo of the Milwaukee Bucks, who is often criticized for taking nearly ten seconds before he shoots. University of Arizona Alumni Gilbert Arenas would wrap the ball around his waist three times before he would shoot, presumably out of superstition. The goal of this project is to investigate some "simple" factors of my own routine, and determine if any of them individually affect how well I shoot free throws.

# 2 Problem Definition, Factors, and Response Variable

"Are there any free throw routine factors that affect my own free throw shooting performance?" This project seeks to identify which of three factors of my own routine, if any, affect my free throw shooting performance. Based on my own experience, the following three were selected for investigation.

- *Time* after ball is received before shooting motion begins (alternatively, time spent looking at rim prior to shot) (Factor A)
- *Angle of stance* relative to free throw line (Factor B)
- Use of *"spring" in legs* for shooting motion (Factor C)

These are the three that I felt best characterized my free throw shooting routine and were reasonably reproducible in an experimental setting.

The quanitiy that will be measured is the number of free throws made in a fixed number of attempts. So, *number of free throws made out of $k$ attempts* is the response variable.

# 3 Experimental Design

The primary goal of the experiment is to *identify* which factors, if any, impact the response. We only seek to estimate the magnitude and direction any individual factor has on the response, and to identify significance to aid in identifying factors that may be worth further investigation. Given these goals, a *Factor Screening Experiment*[1] would suit the goals of the experiment.

### 3.1 $2^3$ Factorial Design (Blocked Replicates), "Factor Screening Experiment"

| | A | B | C |
|---|---|---|---|
| + | 5 counts | ~30 deg angle | "Spring" in legs |
| - | 1 count | 0 deg angle | "Stiff" legs |

Table 1: Factor levels

See Appendix Table 2 for model details. A $2^3$ Factorial experiment was conducted to analyze three *qualitative* factors (factor A could be argued to be approximately *quantitative*) for having a possible effect on my free throw shooting performance. For the analysis, we are treating each of these factors as *fixed*, that is we do not intend to draw inference about some larger population for any particualr factor. There will be further discussion in later sections of the report pertinent to these assumptions.

For this design, the "+" level of each factor corresponds to that of my "natural" free throw shooting routine that I regularly practice. The "-" levels are an alteration of the factor that is different than my "natural"; see Figure 1 comparing the natural shooting form and the "cumbersome" shooting form.

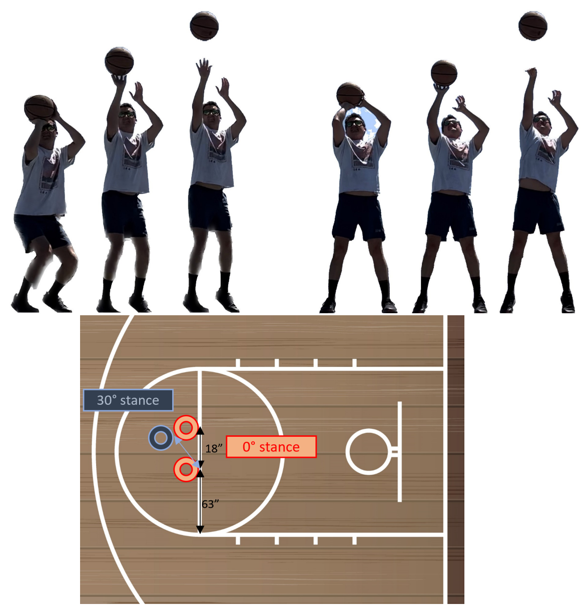



Figure 1: Factor B & C +/+ (left), -/- (right), Stance (below)

The experiment was conducted as a full $2^3$ factorial experiment, a full replicate being performed with 8 total runs ($r_n$) of combinations of factors A/B/C at varying levels. Reference Appendix Table 3 for design matrix, $r_8$ is my "natural" shooting form. Each "run" $r_n\, n = 1, .., 8$ consists of $k$ free throw attempts, and the response recorded is the total number of successful makes in each run. Note that $k$ shots was changed between one of the experimental runs, which will be detailed later in the report. For one experimental run (one replicate), all $r_n\, n = 1, .., 8$ runs are performed *in random order*. Prior to each experiment, the order for $r_n\, n = 1, .., 8$ is randomized without the knowledge of the shooter. Thus this is a *completely randomized experiment*.

Prior to any experimental runs, the foot positions for factor B were marked with small pieces of tape relative to the painted free throw line to aid in repeatability (see Figure 1).

When one experimental run $r_n$ was performed, an assistant would announce which factor combination to be run, the shooter would adjust, and then the assistant would pass the ball to the shooter. Prior to a run beginning, the shooter aligns feet per the factor B level. When the shooter receives the ball, "counts" begin, which the shooter counts by dribbling the basketball (once for each count) while concentrating on the rim, and then shooting with rigid legs or using spring depending upon the factor C level specified for the run. A total of $k$ attempts are taken for the run, and the number made for that run recorded. The assistant then announces the next run factor combination and the experiment continues until all 8 runs (1 replicate) are complete.

Three total replicates of the experimental run were performed at each of three total locations (detailed in next section).

# 4 Data Collection and Analysis

## 4.1 Data Collection

Figure 2: Indoor facility (left), Outdoor II court (right); environmental nuisance factors

Table 4 in Appendix section summarizes the data captured for the $2^3$ factorial design. Three total complete runs of the experiment were completed (each consisting of three replicates), one at each of three different basketball courts. One court was indoors [Indoor], one outdoor with roof cover (somewhat shielded from outdoor environmental factors) [Outdoor I], and one fully outdoor [Outdoor II]. The Indoor and Outdoor I were performed with $k = 5$ shots taken per level. After a cursory look at the data, a significant three factor interaction was produced for both sets of data. There was suspicion that $k = 5$ was not sufficient shots to take per level to properly evaluate accuracy, so the experiment was re-run with $k = 10$ shots per run for the Outdoor II experiment. Note that $k = 5$ shots per run was originally selected with the intent to keep the number of shots taken per experiment reasonable. At $k = 5$ attempts per run, a total of 40 shots are attempted in one full experiment replicate (120 total per location) and is doubled to 80 shots for $k = 10$ (240 total per location).

During the execution of the experiment, several nuisance factors present within each experiment were noted. Some nuisance factors were unique to location. For example, the indoor court had a very "forgiving" rim as compared to the outdoor court; shots hitting the rim at the indoor court bounced much more softly and some shots were made that would likely have not been made bouncing on a "harder" rim. A different basketball was used for the indoor trial than the outdoor trials (although both regulation). Other environmental factors like wind at the outdoor courts, as well as the court type (wood vs. concrete) were noticable differences between experiments. Additionally, as each individual experiment progressed, effects of fatigue were felt as well as possible effects of "getting the feel" for the basket. For this reason, ultimately, it was felt that there was enough run-to-run variation possible that the replicates should be treated as blocks ($\delta_k$ term in Appendix Table 3) in attempts to reduce noise in the experiment attributable to nuisance factors.

## 4.2 Results and Analysis

ANOVA (Indoor)

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 9 | 17.00000000 | 1.88888889 | 2.14 | 0.0970 |
| Error | 14 | 12.33333333 | 0.88095238 | | |
| Corrected Total | 23 | 29.33333333 | | | |

| R-Square | Coeff Var | Root MSE | resp Mean |
|---|---|---|---|
| 0.579545 | 43.31957 | 0.938591 | 2.166667 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| block | 2 | 4.33333333 | 2.16666667 | 2.46 | 0.1215 |
| A | 1 | 1.50000000 | 1.50000000 | 1.70 | 0.2130 |
| B | 1 | 0.66666667 | 0.66666667 | 0.76 | 0.3990 |
| C | 1 | 0.66666667 | 0.66666667 | 0.76 | 0.3990 |
| AB | 1 | 0.16666667 | 0.16666667 | 0.19 | 0.6702 |
| AC | 1 | 1.50000000 | 1.50000000 | 1.70 | 0.2130 |
| BC | 1 | 0.00000000 | 0.00000000 | 0.00 | 1.0000 |
| ABC | 1 | 8.16666667 | 8.16666667 | 9.27 | 0.0087 |

ANOVA (Outdoor I)

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 9 | 15.04166667 | 1.67129630 | 1.24 | 0.3480 |
| Error | 14 | 18.91666667 | 1.35119048 | | |
| Corrected Total | 23 | 33.95833333 | | | |

| R-Square | Coeff Var | Root MSE | resp Mean |
|---|---|---|---|
| 0.442945 | 52.63731 | 1.162407 | 2.208333 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| block | 2 | 3.08333333 | 1.54166667 | 1.14 | 0.3475 |
| A | 1 | 0.37500000 | 0.37500000 | 0.28 | 0.6066 |
| B | 1 | 3.37500000 | 3.37500000 | 2.50 | 0.1363 |
| C | 1 | 0.04166667 | 0.04166667 | 0.03 | 0.8631 |
| AB | 1 | 0.04166667 | 0.04166667 | 0.03 | 0.8631 |
| AC | 1 | 2.04166667 | 2.04166667 | 1.51 | 0.2392 |
| BC | 1 | 1.04166667 | 1.04166667 | 0.77 | 0.3948 |
| ABC | 1 | 5.04166667 | 5.04166667 | 3.73 | 0.0739 |

(a) Indoor & Outdoor I ANOVA, $k = 5$ shots per run

ANOVA (Outdoor II)

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 9 | 22.70833333 | 2.52314815 | 0.95 | 0.5165 |
| Error | 14 | 37.25000000 | 2.66071429 | | |

| R-Square | Coeff Var | Root MSE | resp Mean |
|---|---|---|---|
| 0.378735 | 47.16635 | 1.631170 | 3.458333 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| block | 2 | 4.08333333 | 2.04166667 | 0.77 | 0.4828 |
| A | 1 | 1.04166667 | 1.04166667 | 0.39 | 0.5416 |
| B | 1 | 0.04166667 | 0.04166667 | 0.02 | 0.9022 |
| C | 1 | 12.04166667 | 12.04166667 | 4.53 | 0.0516 |
| AB | 1 | 1.04166667 | 1.04166667 | 0.39 | 0.5416 |
| AC | 1 | 2.04166667 | 2.04166667 | 0.77 | 0.3958 |
| BC | 1 | 2.04166667 | 2.04166667 | 0.77 | 0.3958 |
| ABC | 1 | 0.37500000 | 0.37500000 | 0.14 | 0.7130 |

Parameter Estimates

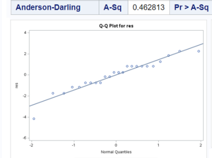
| Variable | Label | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | Intercept | 1 | 2.58333 | 0.86264 | 2.99 | 0.0091 |
| block | | 1 | 0.43750 | 0.39933 | 1.10 | 0.2905 |
| A | A | 1 | 0.20833 | 0.32605 | 0.64 | 0.5325 |
| B | B | 1 | 0.04167 | 0.32605 | 0.13 | 0.9000 |
| C | C | 1 | 0.70833 | 0.32605 | 2.17 | 0.0463 |
| AB | | 1 | -0.20833 | 0.32605 | -0.64 | 0.5325 |
| AC | | 1 | 0.29167 | 0.32605 | 0.89 | 0.3852 |
| BC | | 1 | 0.29167 | 0.32605 | 0.89 | 0.3852 |
| ABC | | 1 | -0.12500 | 0.32605 | -0.38 | 0.7068 |

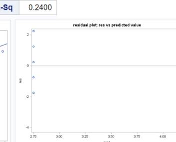
Parameter Estimates

| Variable | Label | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | Intercept | 1 | 3.45833 | 0.30125 | 11.48 | <.0001 |
| C | C | 1 | 0.70833 | 0.30125 | 2.35 | 0.0281 |

Tests for Normality

| Test | Statistic | | p Value | |
|---|---|---|---|---|
| Shapiro-Wilk | W | 0.936579 | Pr < W | 0.1368 |
| Kolmogorov-Smirnov | D | 0.134815 | Pr > D | >0.1500 |
| Cramer-von Mises | W-Sq | 0.06877 | Pr > W-Sq | >0.2500 |
| Anderson-Darling | A-Sq | 0.462813 | Pr > A-Sq | 0.2400 |




(b) Outdoor II ANOVA, $k = 10$ shots (left), full model parameter estimates

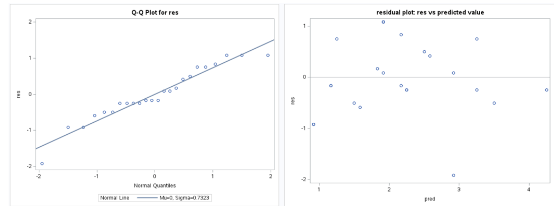
Figure 3: Results

Figure 3a shows ANOVA results for Indoor and Outdoor I experiments, which used $k = 5$ shots per run. No individual factors were significant at an $\alpha = 0.05$ level, nor were any of the individual two-factor interaction effects or the run-to-run block effects. However, an interesting result was found in both Indoor and Outdoor I data. There was a significant interaction at the $\alpha = 0.05$ level for the $ABC$ interaction term, and approximately significance for the same $ABC$ interaction for the Outdoor I data. The interaction term was not of interest to the goals of this experiment and difficult to interpret physically, however it was a possible indicator that the number of shots taken per run was not accurate to analyze due to a small sample size of shots. Effects from a small sample size for this experiment may manifest in the normality diagnostics and constant variance; if the variance of the response does not have adequate fidelity it can impact the validity of our assumptions for the model. Reference Figure 4, our normality and constant variance assumptions for Indoor & Outdoor I data are somewhat borderline in adequacy. For future screening experiments, some additional power analysis would be worthwhile to determine minimum number of shots required per run with more objectivity.

Given time constraints, only one more full experiment was able to be performed, this time with $k = 10$ shots per run. This improved the normality diagnostics (Figure 4). For this

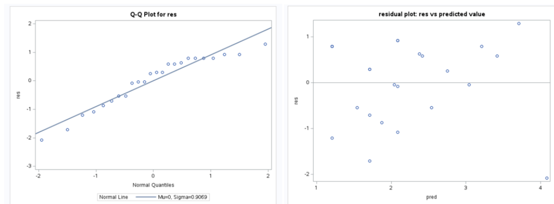
Normality and Constant Variance Indoor

| Tests for Normality | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Shapiro-Wilk | W | 0.945349 | Pr < W | 0.2144 |
| Kolmogorov-Smirnov | D | 0.131688 | Pr > D | >0.1500 |
| Cramer-von Mises | W-Sq | 0.064044 | Pr > W-Sq | >0.2500 |
| Anderson-Darling | A-Sq | 0.428198 | Pr > A-Sq | >0.2500 |




Normality and Constant Variance Outdoor I

| Tests for Normality | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Shapiro-Wilk | W | 0.924349 | Pr < W | 0.0729 |
| Kolmogorov-Smirnov | D | 0.156625 | Pr > D | 0.1296 |
| Cramer-von Mises | W-Sq | 0.117049 | Pr > W-Sq | 0.0648 |
| Anderson-Darling | A-Sq | 0.693365 | Pr > A-Sq | 0.0642 |




Normality and Constant Variance Outdoor II

| Tests for Normality | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Shapiro-Wilk | W | 0.985509 | Pr < W | 0.9723 |
| Kolmogorov-Smirnov | D | 0.111941 | Pr > D | >0.1500 |
| Cramer-von Mises | W-Sq | 0.024512 | Pr > W-Sq | >0.2500 |
| Anderson-Darling | A-Sq | 0.156421 | Pr > A-Sq | >0.2500 |

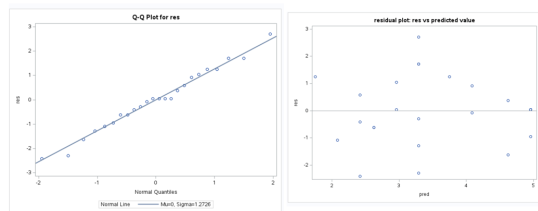



Figure 4: Normality and Constant Variance Diagnostics

experiment, there is approximately significance for the factor C (Spring[+]/Rigid[-]) at the $\alpha = 0.05$ level (Figure 3b), and no significance for any other individual factors, individual interaction terms, or run-to-run block effects. The model was refit to only factor C, and indicated a *positive* change in response (number of free throws made of 10) from the Rigid [-] to Spring [+] factor level. Normality diagnostics for the refit model in Figure 3b are acceptable, so the regression model of least squares estimates applies.

$$y = \beta_0 + \beta_C x_C + \epsilon \quad \epsilon \sim N(0, \sigma^2) \tag{1}$$

with $\beta_0 = 3.458$ and $\beta_C = 0.708$, where $\beta_C$ is half of the factor effect for factor $C$.

# 5 Conclusions and Recommendations

Based on the normality diagnostics and number of shots taken per run, the most reliable data that was collected was the Outdoor II experiment, so inference for this experiment utilize this experiment only.

- Factor C (Spring[+]/Rigid[-]) is approximately significant at the $\alpha = 0.05$ level
- There was not significance at $\alpha = 0.05$ level for run-to-run block effects

- There was not significance at $\alpha = 0.05$ level for any other individual or interaction effects
- Factors are fixed, so inference applies only to these factors and levels

In this screening experiment, Factor C should be of interest to perform some further and more detailed analysis with a different experimental design that provides better focus and fidelity to this factor. Some other similar studies investigating free throw routine for improvements in free throw accuracy found a similar result that the duration of routine (Factor A) does not affect accuracy([2]May, 2011), so it was of a bit of surprise to find any factor that significantly affects my own performance.

We should undertsand the inferences that we are able to draw for this experiment. Our factors are *fixed*, so are conclusions apply only to these three specific factors at the specified levels. Additionally, the inferences for this experiment apply only to the subject (myself), and cannot be generalized to all people. For the goals of this experiment this is okay; this is analysis of my own free throw shooting, and the factor levels are somewhat approximate in nature (none can truly be identically reproduced through each run). We only sought to *identify* factor(s) that affect my free throw shooting, and not necessarily to infer any detailed analysis about them. As with other factorial screening experiments, this study is useful to identify a factor that I may want to perform further analysis to draw more detailed inferences. If I am looking to improve a particular area of my free throw shooting, a focus in optimizing this factor may be useful to me.

# Appendix

The model used for this experiment for the $2^3$ factorial model with blocked replicates is the following.

$$y_{ijkm} = A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk} + \delta_m + \epsilon_{ijkm} \quad \begin{cases} i = 1,2 \\ j = 1,2 \\ k = 1,2 \\ m = 1,2 \end{cases} \tag{2}$$

The hypotheses to be tested will be as follows.

$$H_0 : A_1 = A_2 = 0$$
$$H_1 : \text{at least one of } A_1 \text{ or } A_2 \neq 0$$

And similar hypotheses for factors B and C.

$$H_0 : (AB)_{ij} = 0 \text{ for all } i, j$$
$$H_1 : \text{at least one of } (AB)_{ij} \neq 0$$

And similar hypotheses are extended for interactions $(AC)_{ik}$, $(BC)_{jk}$, and $(ABC)_{ijk}$. Factors are **fixed** (factors were selected and controlled in the experiment), so

$$A_1 + A_2 = 0,\ B_1 + B_2 = 0,\ C_1 + C_2 = 0 \tag{3}$$

$$\sum_{i=1}^{2}(AB)_{ij} = \sum_{j=1}^{2}(AB)_{ij} = 0,\ \sum_{i=1}^{2}(AC)_{ik} = \sum_{k=1}^{2}(AC)_{ik} = 0,\ \sum_{j=1}^{2}(BC)_{jk} = \sum_{k=1}^{2}(BC)_{jk} = 0 \tag{4}$$

$$\sum_{i=1}^{2}(ABC)_{ijk} = \sum_{j=1}^{2}(ABC)_{ijk} = \sum_{k=1}^{2}(ABC)_{ijk} = 0 \tag{5}$$

and $\epsilon_{ijkm} \sim N(0, \sigma^2)$.

Table 2: $2^3$factorial model with blocked replicates, hypotheses, and assumptions

| | A | B | C |
|---|---|---|---|
| $r_1$ | - | - | - |
| $r_2$ | + | - | - |
| $r_3$ | - | + | - |
| $r_4$ | + | + | - |
| $r_5$ | - | - | + |
| $r_6$ | + | - | + |
| $r_7$ | - | + | + |
| $r_8$ | + | + | + |

Table 3: $2^3$ factorial design matrix

| **Indoor** | | | | | |
|---|---|---|---|---|---|
| **Factors** | | | **Replicates** | | |
| **A** | **B** | **C** | **I** | **II** | **III** |
| - | - | - | 2 | 3 | 3 |
| + | - | - | 2 | 1 | 2 |
| - | + | - | 0 | 1 | 3 |
| + | + | - | 3 | 3 | 1 |
| - | - | + | 0 | 1 | 3 |
| + | - | + | 4 | 3 | 4 |
| - | + | + | 2 | 2 | 3 |
| + | + | + | 1 | 2 | 3 |

| **Outdoor I** | | | | | |
|---|---|---|---|---|---|
| **Factors** | | | **Replicates** | | |
| **A** | **B** | **C** | **I** | **II** | **III** |
| - | - | - | 2 | 4 | 3 |
| + | - | - | 2 | 1 | 2 |
| - | + | - | 2 | 2 | 1 |
| + | + | - | 1 | 3 | 3 |
| - | - | + | 1 | 3 | 2 |
| + | - | + | 4 | 2 | 5 |
| - | + | + | 2 | 3 | 0 |
| + | + | + | 0 | 3 | 2 |

| **Outdoor II** | | | | | |
|---|---|---|---|---|---|
| **Factors** | | | **Replicates** | | |
| **A** | **B** | **C** | **I** | **II** | **III** |
| - | - | - | 2 | 5 | 2 |
| + | - | - | 3 | 5 | 1 |
| - | + | - | 1 | 3 | 4 |
| + | + | - | 3 | 2 | 2 |
| - | - | + | 0 | 3 | 6 |
| + | - | + | 4 | 5 | 5 |
| - | + | + | 5 | 3 | 5 |
| + | + | + | 5 | 4 | 5 |

Table 4: Factorial Data

Data analysis using SAS. XLSX files can be interchanged in proc import path. Datasets analyzed are reflected in Table 3.

```
/* Indoor Factorial 2023 APR 15 */

/*SELECT DATA*/
/* alexfactorial20230415.xlsx */
/* alexfactorialoutdoor20230422.xlsx */
/* alexfactorialoutdoor20230425.xlsx */

proc import datafile="/home/.../alexfactorialoutdoor20230425.xlsx"
        dbms=xlsx
        out=alexf1
        replace;
        getnames=yes;
run;

data inter;
        set alexf1;
                AB=A*B;
                AC=A*C;
                BC=B*C;
                ABC=A*BC;
                block=Run;
                resp=of5made;

proc glm data=inter;
 class A B C AB AC BC ABC block;
 model resp=block A B C AB AC BC ABC;
 output out=diag r=res p=pred;
 run;

/* check normality */
proc univariate data=diag normal;
var res;
qqplot res / normal (mu=est sigma=est);
run;

/* check constant variance using graph*/
title 'residual plot: res vs predicted value ';
proc sgplot data=diag;
scatter x=pred y=res;
refline 0;
run;
```

```
proc reg data=inter;
 model resp=block A B C AB AC BC ABC;
run;

/* FIT MODEL TO ONLY FACTOR C */

proc glm data=inter;
 class A B C AB AC BC ABC block;
 model resp=block C;
 output out=diag r=res p=pred;
 run;

/* check normality */
proc univariate data=diag normal;
var res;
qqplot res / normal (mu=est sigma=est);
run;

/* check constant variance using graph*/
title 'residual plot: res vs predicted value ';
proc sgplot data=diag;
scatter x=pred y=res;
refline 0;
run;

proc reg data=inter;
 model resp=block C;
run;
```

# Bibliography


[1] Montgomery, D. C. (2013). Design and Analysis of Experiments (8th ed.). John Wiley & Sons, Inc.

[2] May, Andrew J., "A Comparison of the Effectiveness of Two Free Throw Shooting Methods" (2011). Theses and Dissertations. 2918. https://scholarsarchive.byu.edu/etd/2918

[3] Cannon, Jamaal Edward, "Effects of imagery use in basketball free throw shooting" (2008). Theses Digitization Project. 3354. https://scholarworks.lib.csusb.edu/etd-project/3354