An Experimental Analysis of Time taken by a ball to stop under Various Factors

Submitted to:
Dr. Vandana Sarin Walia
Kirori Mal College, University Enclave, New Delhi-110007

Submitted by:
Gaurav Tomar (2033132)
Vishal Maurya (2033116)
Mayank (2033157)
Ravinandan (2033117)
Shashwat Tiwari(2033160)
Yashi(2033170)

B.Sc. (H) Statistics

Acknowledgement

We would like to extend a special thank you to ma'am Dr.Vandana Sarin Walia and to our principal Dr. Dinesh khattar who gave us a great opportunity to do this great work on the Design of Experiment, which also helped us understand the course contents better and added to our learning about the subject.

We would also like to thank the people who volunteered as subjects for the project, despite their busy plans, helping us in making this project.

Introduction

Design of experiment is the logical construction of an experiment in which degree of uncertainty with which the inference is drawn is well defined. It includes the planning of an experiment obtaining the relevant information from it regarding the statistical hypothesis under consideration and making the statistical analysis of the data. In simpler words, it is deciding how observations/measurement should be taken to answer a particular question in a valid economic and efficient manner, the design and final analysis go hand in hand.

Having century old humble origins in agriculture[i] , today it finds application in pharmaceutical[ii] , machine learning and big data[iii] , nanotechnology[iv] , social sciences, sciences and various other fields.

Personally, I understand DOE as a useful tool that helps one seek multiple goals with minimum resource utilization, while creating a robust product/process, using the sequential optimization methodology.

Objective

The immediate objective of this project is to implement the common statistical techniques under “Design of Experiment” by analysing the time taken by a ball to stop when dropped from a certain height under influence of certain effects.

Key advantage of method being the experimental effort is kept to a minimum while still achieving the immediate goal.[1] Another goal was to gain a better understanding of the process/method.

Using data collected manually, this project estimates the effect of following factors on the time taken by a ball to stop when dropped from a certain height:

1. Effect of size of ball
2. Effect of type of surface on which ball is dropped
3. Effect of height from which ball is dropped

Design

A factorial experiment (also known as a complex experiment) is an experiment in which we can study a number of factors simultaneously in a single experiment. In a factorial experiment we study not only the main effects of factors but also their interaction effects. These kinds of experiments are economically more viable than the single factor experiments (both cost efficient and time wise viable).

Factorial Completely Randomized Design is an experimental design in which the treatment is formed by a combination of the levels of several factors. Factorial treatment structures are useful for examining the effects of two or more factors on a response y, whether or not interaction exists. The treatment design of the multi-factor experiment was differentiated based on the level of importance and restrictions on randomization of the levels of each factor making up the treatment.

Since every observation doesn't depend on how an individual acts, so there can't be homogeneous blocks. And RBD is not appropriate.

Hence, CRD is appropriate only for experiments with homogeneous experimental units, such as laboratory experiments, where environmental effects are relatively easy to control. For field experiments, where there is generally large variation among experimental plots in such environmental factors as soil, the CRD is rarely used.

This is a $\mathbf{2^3}$ **factorial** experiment, i.e. it has 3 factors namely,

A: denoting size of ball, B: denoting surface on which ball is dropped ,
C denoting height from which the ball is dropped,
Which are at two levels each, 0 and 1, where,

0: indicates the presence of the factor
1: indicates the absence of the factor

Model: $y_{ijkl} = \mu + a_i + b_j + c_k + ab_{ij} + bc_{jk} + ac_{ik} + abc_{ijk} + \varepsilon_{ijkl}$

Where the symbols have their usual meaning and $\varepsilon \sim NID(0, \sigma^2)$

Hypothesis:

H_{0i}: There is no significant effect due to the Main Effect i ; i = A,B,C
H_{1i}: There is a significant effect due to the Main Effect i ; i = A,B,C

H_{0ij}: There is no significant effect due to the 2FIE ij ; i,j = A,B,C
H_{1ij}: There is a significant effect due to the 2FIE ij ; i,j = A,B,C

H_{07}: There is no significant effect due to the 3FIE ABC
H_{17}: There is a significant effect due to the 3FIE ABC

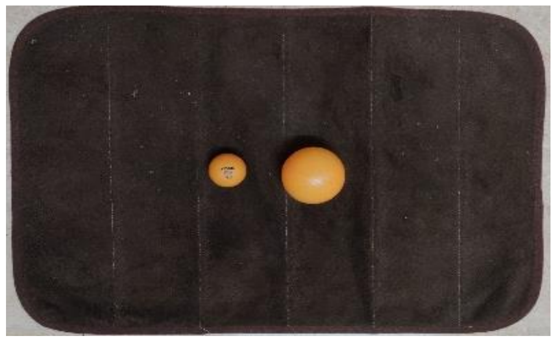
Measurement of Variables

The effects in our experiment are as follows:

A: Size of the ball

a_0: Dropping a small ball

a_1: Dropping a big ball



B: Type of surface on which ball is dropped

b_0: Dropping on tile

b_1: Dropping on carpet

C: Height from which ball is dropped

c_0: Dropping from 1.5 m

c_1: Dropping from 2 m

Hence,

a_0	Small ball
a_1	Big ball
b_0	Dropped on tile
b_1	Dropped on carpet
c_0	Dropped from 1.5 m
c_1	Dropped from 2 m

<u>Observed data is recorded in seconds</u>

<u>Data</u>

	$a_0b_0c_0$	$a_1b_0c_0$	$a_0b_1c_0$	$a_1b_1c_0$	$a_0b_0c_1$	$a_1b_0c_1$	$a_0b_1c_1$	$a_1b_1c_1$
Subject No.	**1**	**a**	**b**	**ab**	**c**	**ac**	**bc**	**abc**
1	11.05	5.72	2.74	2.81	12.96	7.36	2.78	3.07
2	11.14	5.81	2.9	2.63	12.81	7.05	2.9	3.65
3	11.75	5.88	2.51	2.41	13.01	7.79	2.63	3.23
4	11.26	5.61	2.9	2.29	12.91	7.17	2.35	3.5
5	10.82	5.45	2.34	2.69	12.24	6.96	2.58	3.11
6	11.33	5.83	2.76	2.71	13.55	6.83	2.36	3.58
7	11.12	5.55	2.56	2.79	12.23	7.22	2.64	3.22
8	11.19	5.73	2.33	2.56	12.91	6.9	2.78	3.4
9	11.2	5.98	2.8	2.44	12.89	7.12	2.81	3.45
10	10.97	5.43	2.65	2.68	13.01	7.29	2.7	3.53

Limitations

1. Since no pre-survey or study was conducted, impact level of effects and their interactions was unknown, hence a full factorial experiment had to be conducted.

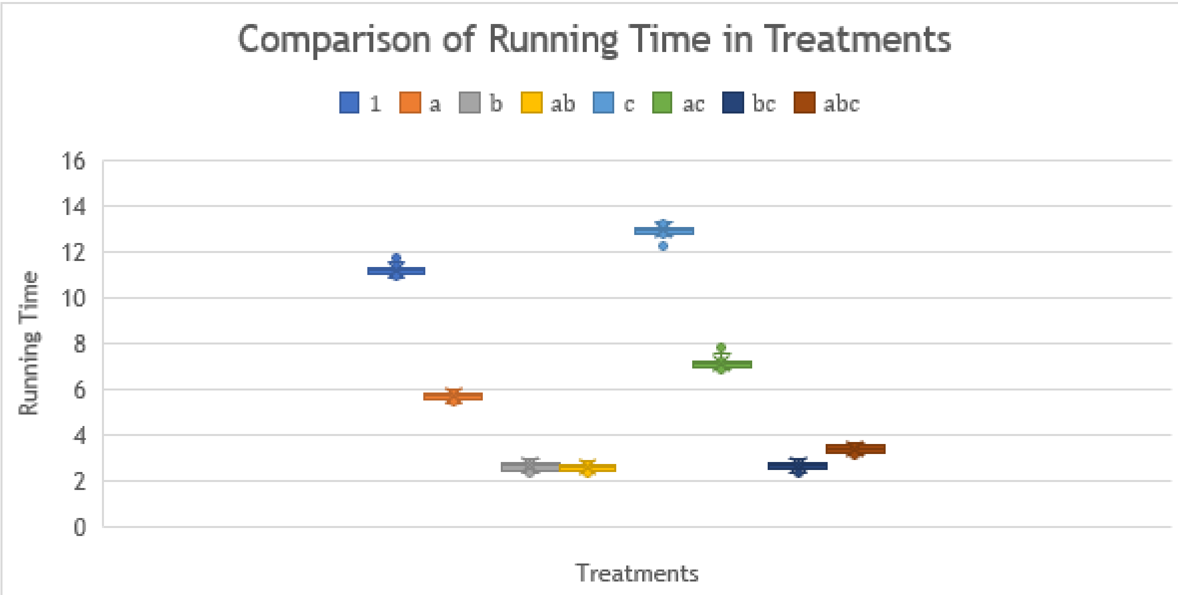
2. Other factors like air also affects the observations but it was not practically possible for us to measure.
3. The time lapse cannot be avoided i.e., there is always a difference between the actual time when ball is dropped and time started in stop watch, however we tried to minimize it.

The following things were kept in mind:

- While bouncing a ball, any obstruction may impact the observation. So, we have taken observation again to avoid it.
- We tried to minimize external forces acting on a ball as this can make ball to go side and affect the time.

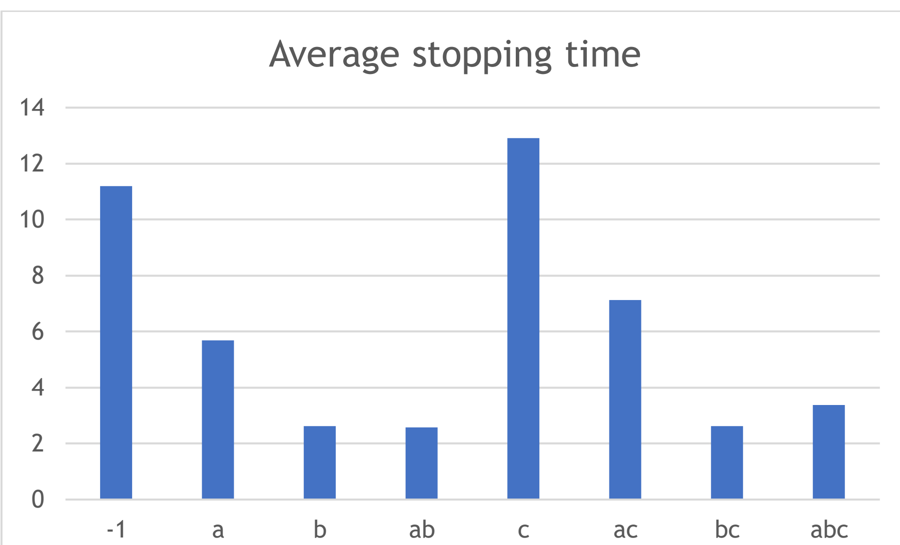
Analysis

Boxplot



From the box-plot, we can conclude that data with each factor combination is very condensed. This is because all the observation with a factor combination is doesn’t affected by individual dropping the ball.

Bar-plot:



- From the bar-plot, we can conclude that there is a considerable difference between averages of different factor combinations.
- We can observe that a small ball dropped on tile from 2m is most likely to be highest stopping time among the all factor combinations.

Anova

For the model

$$y_{ijkl} = \mu + a_i + b_j + c_k + ab_{ij} + bc_{jk} + ac_{ik} + abc_{ijk} + \varepsilon_{ijkl}$$

Where the symbols have their usual meaning and $\varepsilon \sim NID(0, \sigma^2)$

Tests of Between-Subjects Effects

Dependent Variable: Y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1161.472[a]	7	165.925	2843.893	.000
Intercept	2901.641	1	2901.641	49733.172	.000
A	137.655	1	137.655	2359.363	.000
B	820.865	1	820.865	14069.356	.000
C	19.169	1	19.169	328.547	.000
A * B	175.232	1	175.232	3003.419	.000
A * C	.406	1	.406	6.961	.010
B * C	6.974	1	6.974	119.529	.000
A * B * C	1.171	1	1.171	20.075	.000
Error	4.201	72	.058		
Total	4067.313	80			
Corrected Total	1165.673	79			

a. R Squared = .996 (Adjusted R Squared = .996)

Comparisons of significant effects

A

Dependent Variable: Y

A	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
a0	7.334	.038	7.258	7.410
a1	4.711	.038	4.635	4.787

Small ball is most likely to takes more time to stop.

3. B

Dependent Variable: Y

B	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
b0	9.226	.038	9.150	9.302
b1	2.819	.038	2.743	2.895

Ball dropped on tile is most likely to takes more time to stop.

4. C

Dependent Variable: Y

C	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
c0	5.533	.038	5.457	5.609
c1	6.512	.038	6.436	6.588

Ball dropped from height of 2m is most likely to take more time to stop.

5. A * B

Dependent Variable: Y

A	B	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
a0	b0	12.018	.054	11.910	12.125
	b1	2.651	.054	2.543	2.759
a1	b0	6.434	.054	6.326	6.542
	b1	2.987	.054	2.880	3.095

Small ball dropped on tile will take more time to stop among all the factor combinations of A and B.

6. A * C

Dependent Variable: Y

A	C	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
a0	c0	6.916	.054	6.808	7.024
	c1	7.753	.054	7.645	7.860
a1	c0	4.150	.054	4.042	4.258
	c1	5.271	.054	5.164	5.379

Big ball and height of 1.5m take less time to stop or small ball and height of 2m takes more time.

7. B * C

Dependent Variable: Y

B	C	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
b0	c0	8.441	.054	8.333	8.549
	c1	10.011	.054	9.903	10.118
b1	c0	2.625	.054	2.517	2.733
	c1	3.013	.054	2.906	3.121

Ball dropped on tile from 2 m will take more time to stop among all the factor combinations of B and C.

8. A * B * C

Dependent Variable: Y

A	B	C	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
a0	b0	c0	11.183	.076	11.031	11.335
		c1	12.852	.076	12.700	13.004
	b1	c0	2.649	.076	2.497	2.801
		c1	2.653	.076	2.501	2.805
a1	b0	c0	5.699	.076	5.547	5.851
		c1	7.169	.076	7.017	7.321
	b1	c0	2.601	.076	2.449	2.753
		c1	3.374	.076	3.222	3.526

Small ball dropped on tile from 2 m will take more time to stop among all the factor combinations of B and C.

Conclusion

Based on the ANOVA table, at a significance level of α=0.05 and we can conclude that of all the effects and their interactions are significant. That means all the factors i.e. size of ball(A), surface on which ball is dropped(B) and height from which we are dropping(C) are affecting this experiment predominantly. Coupling the results of ANOVA with the boxplot and bar plot, it is easy to observe that every factor affects the time significantly.

The results also make sense based on the practical knowledge we have.

Based on this experiment, we can say :

Bigger will be the size of ball it is less likely to bounce.

There will be a better bounce in tiles rather carpet.

As there is increase in the height of the ball dropped there will be an increase in the height of the bounce.

Bibliography

[i] 1.1 - A Quick History of the Design of Experiments (DOE) | STAT 503 (psu.edu)

[ii] Design of experiments (DoE) in pharmaceutical development - PubMed (nih.gov)

[iv] Review Article (chesci.com)

Other resources used

(1) *Fundamentals of Applied Statistics (S.C Gupta & V.K Kapoor)*
(2) *Design of Analysis Of Experiments (D. C. Montogomery)*
(3) *Classroom Notes (Dr. Vandana Sarin Walia)*