Project 1 - Helicopters!

Jaskaran Singh, Michael Sarkis, Jack McCarthy

3/7/2022

Introduction

The aim of this project is to optimize certain design parameters for the construction helicopter toys. The ideal result is a helicopter design that maximizes hang-time with minimal variance in the times across drops. However, as the budget is limited for the testing of various designs, we cannot perform more than 100 drops across all candidate designs. It is important that we perform these test drops in as economical a fashion as possible–that is, we would like to extract the most information possible about the effects of specific design decisions from each trial run as we can.

To approach the given design problem, we initially identified relevant design parameters which appeared to affect the hang time, utilizing 10 runs out of 100. Further, we hypothesized that higher order interactions between these design factors would be less relevant and designed a fractional factorial experiment to estimate the main effects and 2 factor interactions for the design parameters of interest.

After careful experimental design consideration, we collected data for helicopter hang time across the replicates for various combinations of design factors. This data was analyzed to ascertain the main effects and 2 factor interactions which further helped to model location and dispersion for hang time based on these fractional factorial effects.

The location and dispersion models for hang time could further be used to optimize the design choice to reach the overall objective of maximum hang time with minimal variance.

Design

The overall design choice was a combination of sequential runs and fractional factorial design. 10 sequential runs were performed to understand how various design choices such as wing length, wing width, stabilizer length, stabilizer width, stabilizer fold, adding clippers to stabilizer etc. affect the hang time. A couple of radically new helicopter designs were also tried to explore possibility of a completely different starting point for optimal design of helicopters. Based on these preliminary results, we finalized on the seemingly relevant factors for helicopter design which are mentioned below:

Factor	**Description**	**Low (-)**	**High (+)**
A	Total size	6.5”	11.0”
B	Blade length	4.75”	5.75”
C	Blade width	2.1”	3.0”
D	Stabilizer length	2.75”	4.75”
E	Stabilizer width	0.5”.	1.25”

For convenience we are referring to the bottom section of the helicopter that extends downward beneath the blades as the “stabilizer”. We also noted during our initial tests that taping the bottom of the stabilizer is universally beneficial to the hang-time and stability of the helicopter. We will therefore being applying tape in all designs.

The factors of interest mentioned above are blade length, blade width, stabilizer length and stabilizer width with each level of a factor proportional to two different sized helicopters. We finalized on these levels based on our observations on hang time from initial 10 sequential runs.

Since, the overall objective of maximizing hang time involved minimizing variation as well, we restricted our choice to 5 factors giving us more degrees of freedom for replicates and thus helping us to get accurate estimate for variation in hang time. Further, we assumed that higher order interactions between these factors would be minimal and based on this assumption we designed a fractional factorial design to estimate main effects and 2 factor interactions. We selected a 2_V^{5-1} fractional factorial experiment to analyze the effects of these fractional factors on the hang-time as resolution V design has strongly clear main effects and clear two-factor interactions. The generator for this experiment is $I = ABCDE$. This resulted in 16 distinct factor combinations with 5 replicates for each combination, totaling 80 drops (80 degrees of freedom). After consuming 10 initial sequential runs and 80 runs for fractional factorial design, the remaining 10 runs are used to re-validate the top two design choices based on results from fractional factorial experiment.

Analysis

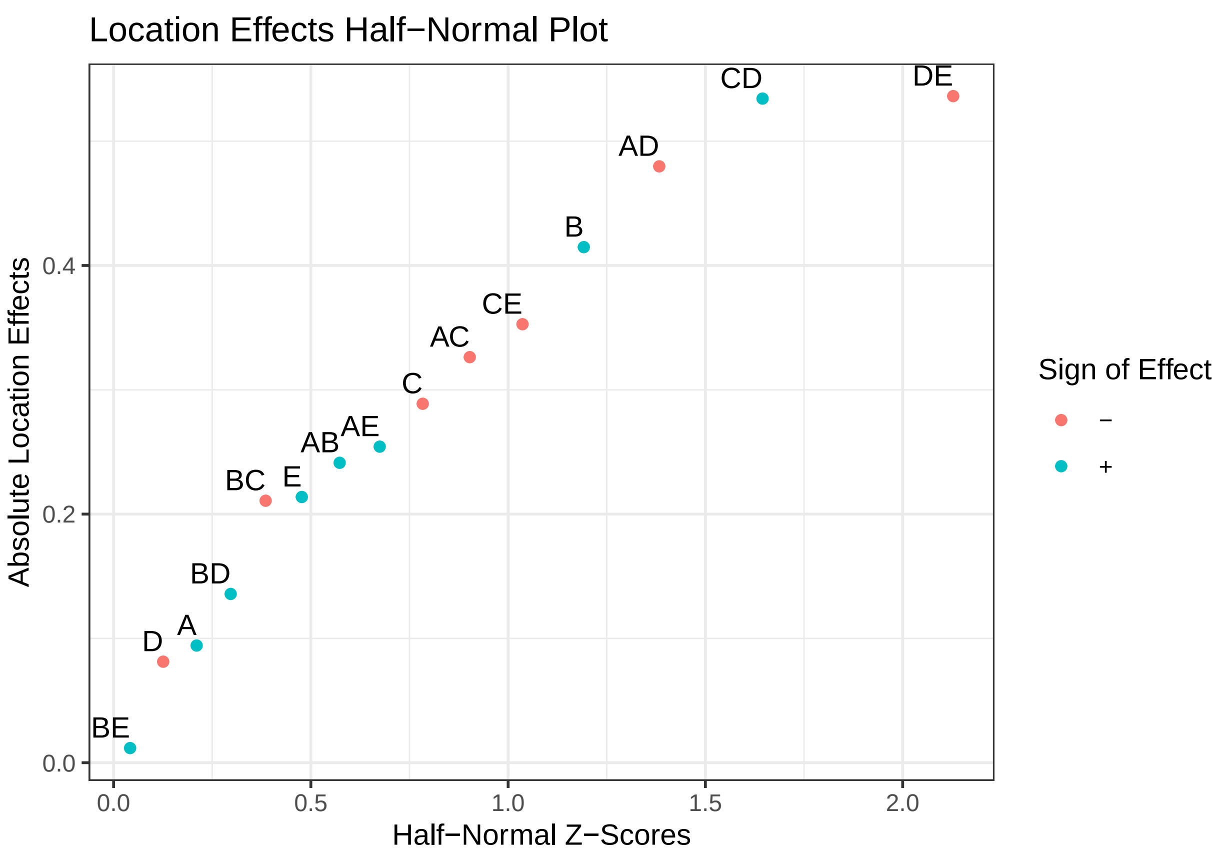
The 5 factor combinations and collected data for hang time are given in the following table:

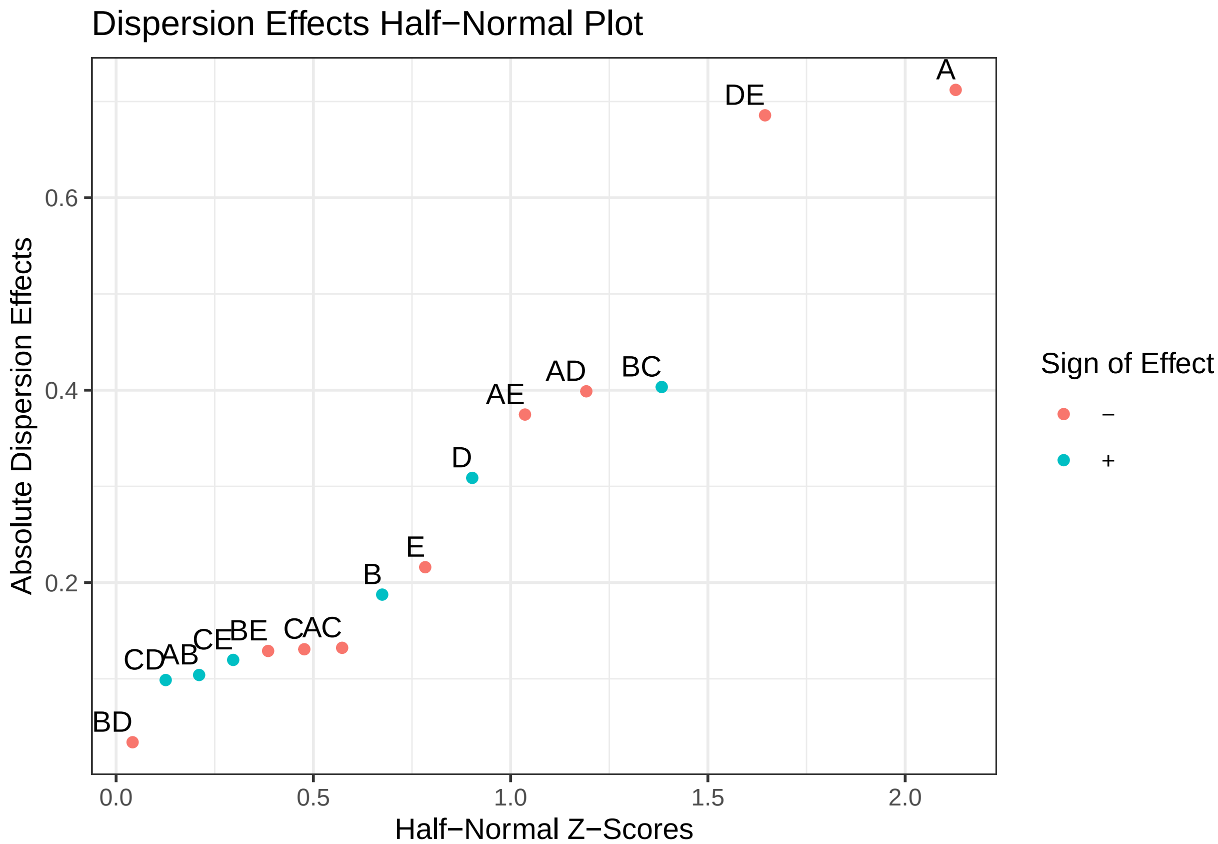
A	**B**	**C**	**D**	**E**	1	2	3	4	5
-	-	-	-	+	5.05	3.61	4.20	4.20	4.72
-	-	-	+	-	2.84	3.94	3.30	4.45	4.39
-	-	+	-	-	3.29	3.35	3.68	3.68	3.93
-	-	+	+	+	4.12	3.67	4.66	3.36	4.08
-	+	-	-	-	3.35	3.74	3.55	3.93	4.21
-	+	-	+	+	4.13	3.95	3.54	4.72	4.06
-	+	+	-	+	3.21	3.54	3.22	3.29	4.52
-	+	+	+	-	4.39	5.31	4.32	4.85	6.16
+	-	-	-	-	4.13	4.27	4.13	3.35	3.74
+	-	-	+	+	3.80	3.68	3.15	3.34	3.54
+	-	+	-	+	4.08	3.87	3.55	4.32	4.07
+	-	+	+	-	3.69	3.36	4.21	3.42	3.88
+	+	-	-	+	5.83	5.58	5.96	6.16	6.53
+	+	-	+	-	3.84	3.62	4.72	4.66	4.26
+	+	+	-	-	3.66	3.42	4.27	3.47	3.36
+	+	+	+	+	3.61	3.68	4.14	4.19	3.74

The above dataset is analysed to first calculate main effects and 2 factor interactions. The location and dispersion effects for factor combinations is shown below:

These location and dispersion effects are tested for significance using half-normal plots and Lenth's method. The plots given below show fractional factorial effects for location and dispersion on half-normal plots.

label	loc.eff	disp.eff
A	0.09425	-0.7120771
B	0.41475	0.1874868
C	-0.28875	-0.1306772
D	-0.08125	0.3087923
E	0.21375	-0.2159567
AB	0.24125	0.1039309
AC	-0.32625	-0.1321951
AD	-0.47975	-0.3987784
AE	0.25425	-0.3746291
BC	-0.21075	0.4032653
BD	0.13575	-0.0340412
BE	0.01175	-0.1289267
CD	0.53425	0.0986884
CE	-0.35275	0.1195921
DE	-0.53625	-0.6856557





These effects are further tested for significance using Lenth's method. Testing for significance via Lenth's method, we find that there are no significant effects for location, but **A** and **DE** are significant for dispersion as we originally surmized from the half-normal plots.

One limitation of the above analysis is that the reading for hang time were taken in open space near a staircase which might have had an impact on hang time and its variance leading to false insignificance of fractional effects under Lenth's method. However, looking at the collected data, the top three optimal designs show very little variation in hang time readings. So, it might be a reasonable assumption that though wind might have had an impact on mean location effect for hang time, its impact on variance of hang time could be minimal and since we are interested in relative comparison of designs for hang time, the effect of wind could be disregarded to some extent.

Interpreting the factor A and DE, the size of helicopter is represented by A, stabilizer length represented by D and stabilizer width represented by E, DE represents interaction of stabilizer length and width.

Optimization

The identified significant fractional effects could be used to find the optimal design. As we observed from the above analysis, there weren't any significant fractional factorial effects for location of hang time. The factors A and DE were significant for dispersion of hang time. The optimal choice was to keep A (total size) at 11.0", D (stabilizer length) at 2.75", E (stabilizer width) at 1.25" which maximizes the observed hang time and minimizes variance. Since there were not conflicting factors affecting location of hang time as compared to its variance, this optimization choice for design was quite straightforward.

This design choice addresses the given problem as keeping factors at optimized levels, reduced variance in hang time to minimum while ensuring maximum hang time.

After selecting the optimal design choice, validation was performed using top two design choices, namely, ABCDE=++−+ and ABCDE=| -+++-. Since the final evaluation is based on minimum hang time, it was of utmost importance to evaluate final two designs based on this metric as a validation criteria. The 5 runs each for top two designs were taken which ultimately corroborated our belief in the optimized design we had chosen ABCDE=++−+.

The readings taken in validation runs for top two designs are as follows:

A	**B**	**C**	**D**	**E**	1	2	3	4	5
-	+	+	+	-	4.77	5.39	5.34	4.69	6.24
+	+	-	-	+	5.91	5.70	5.82	6.21	6.56

Conclusion

We recommended design ABCDE=++−+ which ensures maximum hang time with minimal variance. This can be interpreted as Total size=11", Blade length=5.75", Blade width=2.1", Stabilizer length=2.75" and Stabilizer width=1.25"

Appendix

```
# Location, dispersion effect functions
main.loc <- function(df, fac, resp) {
  # for column selection
  fac  <- sym(fac)
  resp <- enquo(resp)

  # calculate main effect
  df %>%
    group_by_at(1:5) %>%
    summarise(y.bar=mean(!!resp)) %>%
    group_by(!!fac) %>%
    summarise(mean.y.bar=mean(y.bar)) %>%
    mutate(to.add=!!fac*mean.y.bar) %>%
    pull(to.add) %>% sum()
}

main.disp <- function(df, fac, resp) {
  # for column selection
  fac  <- sym(fac)
  resp <- enquo(resp)

  # calculate main effect
  df %>%
    group_by_at(1:5) %>%
    summarise(lns.2=log(var(!!resp))) %>%
    group_by(!!fac) %>%
    summarise(mean.lns.2=mean(lns.2)) %>%
    mutate(to.add=!!fac*mean.lns.2) %>%
    pull(to.add) %>% sum()
}
```

```
two.fac.loc <- function(df, fac, resp) {
  facs <- strsplit(fac, "")[[1]]
  fac1 <- sym(facs[1])
  fac2 <- sym(facs[2])
  resp <- enquo(resp)

  # calculate two-factor location
  eff.1 <- df %>%
    filter(!!fac2>0) %>%
    main.loc(., facs[1], !!resp)
  eff.2 <- df %>%
    filter(!!fac2<0) %>%
    main.loc(., facs[1], !!resp)
  0.5 * (eff.1 - eff.2)
}

two.fac.disp <- function(df, fac, resp) {
  facs <- strsplit(fac, "")[[1]]
  fac1 <- sym(facs[1])
  fac2 <- sym(facs[2])
  resp <- enquo(resp)

  # calculate two-factor dispersion
  eff.1 <- df %>%
    filter(!!fac2>0) %>%
    main.disp(., facs[1], !!resp)
  eff.2 <- df %>%
    filter(!!fac2<0) %>%
    main.disp(., facs[1], !!resp)
  0.5 * (eff.1 - eff.2)
}

# Effects tables code
eff.tbl %>%
  mutate(sgn=as.factor(sign(loc.eff)),
         loc.eff=abs(loc.eff)) %>%
  arrange(loc.eff) %>%
  mutate(i=1:nrow(.)) %>%
  mutate(z=qnorm(((i - 0.5)/nrow(.) + 1)/2)) %>%
  ggplot(aes(x=z, y=loc.eff)) +
    geom_point(aes(color=sgn)) +
    geom_text(aes(label=label),
              hjust=1, vjust=-0.5) +
    labs(x="Half-Normal Z-Scores",
         y="Absolute Location Effects",
         title="Location Effects Half-Normal Plot",
         color="Sign of Effect") +
    scale_color_discrete(labels=c("-", "+")) +
    theme_bw()
```

```
eff.tbl %>%
  mutate(sgn=as.factor(sign(disp.eff)),
         disp.eff=abs(disp.eff)) %>%
```

```r
  arrange(disp.eff) %>%
  mutate(i=1:nrow(.)) %>%
  mutate(z=qnorm(((i - 0.5)/nrow(.) + 1)/2)) %>%
  ggplot(aes(x=z, y=disp.eff)) +
    geom_point(aes(color=sgn)) +
    geom_text(aes(label=label),
              hjust=1, vjust=-0.5) +
    labs(x="Half-Normal Z-Scores",
         y="Absolute Dispersion Effects",
         title="Dispersion Effects Half-Normal Plot",
         color="Sign of Effect") +
    scale_color_discrete(labels=c("-", "+")) +
    theme_bw()
```

```r
# Lenth's method code
loc.all <- abs(eff.tbl$loc.eff)
cutoff <- 3.75 * median(loc.all)
int.sub <- loc.all[loc.all <= cutoff]
pse <- 1.5 * median(int.sub)

lenth.med <- eff.tbl %>%
  mutate(loc.abs=abs(loc.eff),
         disp.abs=abs(disp.eff)) %>%
  mutate(loc.cut=3.75*median(loc.abs),
         disp.cut=3.75*median(disp.abs)) %>%
  mutate(loc.inc=ifelse(loc.abs<=loc.cut, T, F),
         disp.inc=ifelse(disp.abs<=disp.cut, T, F)) %>%
  filter(loc.inc, disp.inc) %>%
  summarise(loc.pse=1.5*median(loc.abs),
            disp.pse=1.5*median(disp.abs)) %>%
  mutate(ier=2.16)

lenth.tbl <- bind_cols(eff.tbl, lenth.med) %>%
  mutate(loc.sig=abs(loc.eff)/loc.pse > ier,
         disp.sig=abs(disp.eff)/disp.pse > ier) %>%
  mutate(label=labels) %>%
  select(label, loc.sig, disp.sig)

lenth.tbl %>% kable(col.names=c("Factors",
                                "Loc. Significant?",
                                "Disp. Significant?"),
                    align=c('l', 'c', 'c')) %>%
  kable_styling(position="center") %>%
  row_spec(0, bold=T)
```