Exam topics:

The exam will cover the topics from March 7 to April 8 as outlined on the weekly content section the course webpage.

Review questions:

1. Write a $2^3$ design in blocks of size 2 such that no main effect is confounded with block differences.

2. Write a $2^4$ design in four blocks such that the main effects are not confounded with block differences.

3. In a research study two different treatments were applied to the two eyes of patients by flipping a fair coin. But the treatment applied to one eye had an influence on the response in the other eye. Is this an appropriate experimental design to study compare the effectiveness of the treatments?

4. Describe the design in the following scenarios:

    a) Suppose that four medical treatments (A, B, C, D) were to be compared at four clinical research sites. But each clinical research site could only collect data on three treatments. The consulting statistician suggested using hospital as a blocking variable. The design is shown in the following table:

| Hospital | Treatments | | |
|---|---|---|---|
| 1 | A | B | C |
| 2 | A | B | D |
| 3 | A | C | D |
| 4 | B | C | D |

    b) Fifteen judges rated two randomly allocated brands of beer, A and B, according to taste (scale: 1 to 10).

    c) Twenty judges each rated two brands of beer, A and B, according to taste (scale: 1 to 10).

    d) You have 16 benches in a greenhouse, that differ in light and temperature such that each bench has different conditions associated with 4 distinct light and 4 temperature levels: 25%, 50%, 75% and 100% of full sun; 15, 20, 25, 30 C. 4 fertilizer treatments are applied to the benches (75, 150,

225 and 300 mM nitrogen), randomized in such a way that every fertilizer treatment is tested under all 4 light levels, and all 4 temperature levels.

5. The following experimental results were obtained from a welding study.

A: open circuit voltage
B: slope
C: electrode melt-off rate (ipm)
D: electrode diameter (in)
E: electrode extension (in)

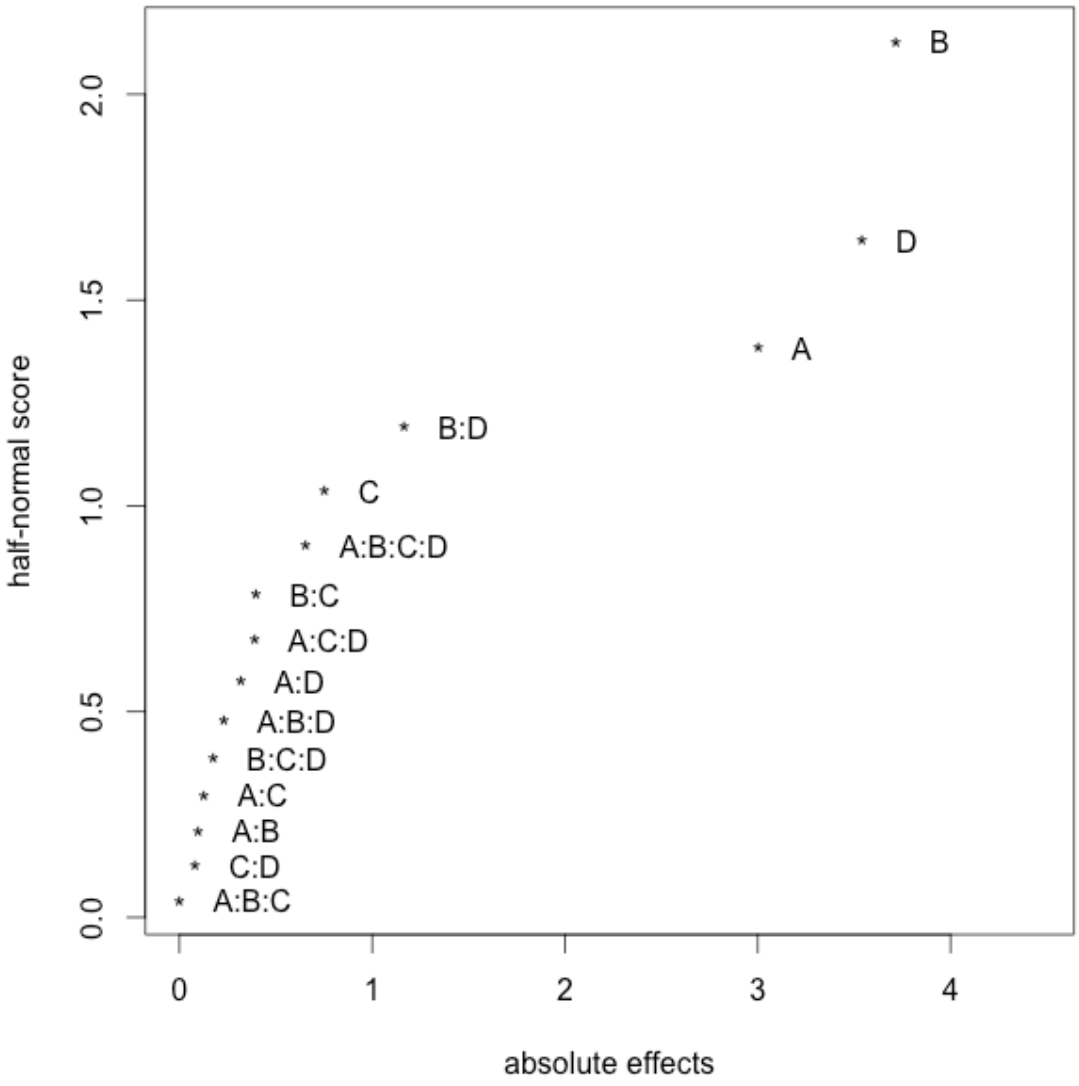
| | runs | A | B | C | D | E | y1 |
|---|---|---|---|---|---|---|---|
| 1 | 12 | -1 | -1 | -1 | -1 | 1 | 23.43 |
| 2 | 1 | 1 | -1 | -1 | -1 | -1 | 25.70 |
| 3 | 2 | -1 | 1 | -1 | -1 | -1 | 27.75 |
| 4 | 6 | 1 | 1 | -1 | -1 | 1 | 31.60 |
| 5 | 15 | -1 | -1 | 1 | -1 | -1 | 23.57 |
| 6 | 8 | 1 | -1 | 1 | -1 | 1 | 27.68 |
| 7 | 7 | -1 | 1 | 1 | -1 | 1 | 28.76 |
| 8 | 4 | 1 | 1 | 1 | -1 | -1 | 31.82 |
| 9 | 11 | -1 | -1 | -1 | 1 | -1 | 27.09 |
| 10 | 14 | 1 | -1 | -1 | 1 | 1 | 31.28 |
| 11 | 3 | -1 | 1 | -1 | 1 | 1 | 31.20 |
| 12 | 16 | 1 | 1 | -1 | 1 | -1 | 33.42 |
| 13 | 13 | -1 | -1 | 1 | 1 | 1 | 29.51 |
| 14 | 10 | 1 | -1 | 1 | 1 | -1 | 31.35 |
| 15 | 5 | -1 | 1 | 1 | 1 | -1 | 31.16 |
| 16 | 9 | 1 | 1 | 1 | 1 | 1 | 33.65 |

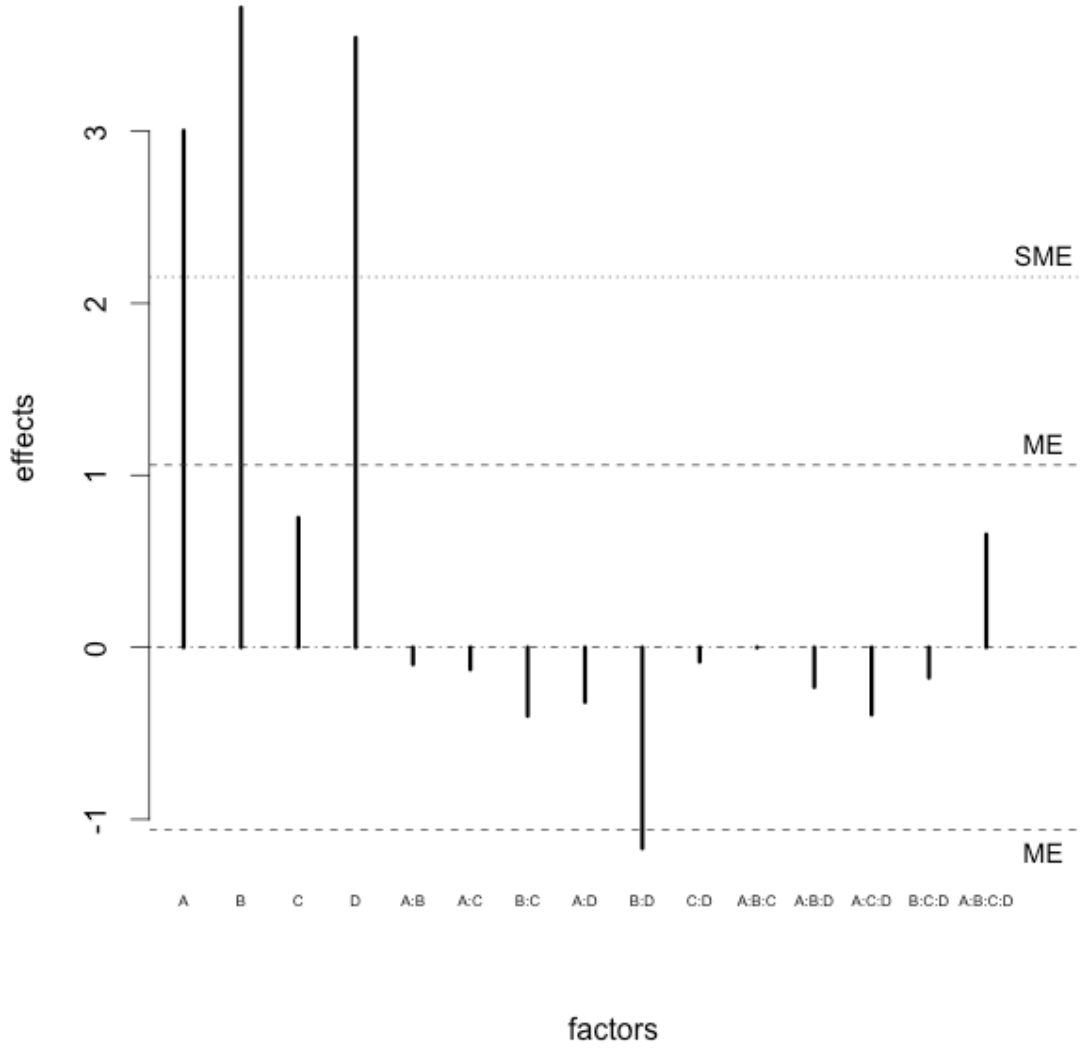
The data were analysed in R.

```
> fact1 <- lm(y1~A*B*C*D,data=prb0609)
> summary(fact1)

Call:
lm.default(formula = y1 ~ A * B * C * D, data = prb0609)

Residuals:
ALL 16 residuals are 0: no residual degrees of freedom!

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 29.310625         NA      NA       NA
A            1.501875         NA      NA       NA
B            1.859375         NA      NA       NA
C            0.376875         NA      NA       NA
D            1.771875         NA      NA       NA
A:B         -0.049375         NA      NA       NA
A:C         -0.064375         NA      NA       NA
B:C         -0.199375         NA      NA       NA
A:D         -0.159375         NA      NA       NA
B:D         -0.584375         NA      NA       NA
C:D         -0.041875         NA      NA       NA
A:B:C       -0.000625         NA      NA       NA
A:B:D       -0.115625         NA      NA       NA
A:C:D       -0.195625         NA      NA       NA
B:C:D       -0.088125         NA      NA       NA
A:B:C:D      0.328125         NA      NA       NA

Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared:      1,   Adjusted R-squared:    NaN
F-statistic:   NaN on 15 and 0 DF,  p-value: NA
```

```
> fact2 <- aov(y1~A*B*C*D,data=prb0609)
```

```
> DanielPlot(fact2,half=T)
> LenthPlot(fact2,cex.fac=0.5)
   alpha      PSE       ME      SME
0.050000 0.412500 1.060365 2.152694
```






a) What type of experimental design was implemented? Explain

b) Is there any aliasing present in the design? If aliasing is present then specify the defining relation and all the aliasing relations.

c) Is this design replicated? If it is replicated then estimate the standard error of the effects.

d) Which factors are significant according to Lenth's method? Are the same factors statistically significant at the 5% level if you use the half-normal plot instead of Lenth's method? Explain.

e) What is estimated effect of the interaction between B and D?

6. The following experiment investigated two factors, pulp preparation method and temperature, on the tensile strength of paper. Temperature was to be set at four levels, and there were three preparation methods. It was desired to run three replicates, but only 12 runs could be made per day. One replicate was run on each of the three days.

   On each day, three batches of pulp were prepared by the three different methods. Each of the three batches was subdivided into four equal parts, and processed at a different temperature. The data are given in the table below.

| | | Day 1 | | | Day 2 | | | Day 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Prep. method | | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| Temp | 1 | 30 | 34 | 29 | 28 | 31 | 31 | 31 | 35 | 32 |
| | 2 | 35 | 41 | 26 | 32 | 36 | 30 | 37 | 40 | 34 |
| | 3 | 37 | 38 | 33 | 40 | 42 | 32 | 41 | 39 | 39 |
| | 4 | 36 | 42 | 36 | 41 | 40 | 40 | 40 | 44 | 45 |

The data were analysed in R:

```
> temp <- factor(c(rep(1,9),rep(2,9),rep(3,9),rep(4,9)))
> prep <- factor(c(rep(c(1,2,3),12)))
> day <- factor(rep(c(rep(1,3),rep(2,3),rep(3,3)),4))
> strength <-
c(30,34,29,28,31,31,31,35,32,35,41,26,32,36,30,37,40,34,37,38,33,40,42,
32,41,39,39,36,42,36,41,40,40,40,44,45)
>
> tensdata <- data.frame(day,prep,temp,strength)
>
```

```
> model1 <- aov(strength~temp*prep,data=tensdata)
> summary(model1)
            Df Sum Sq Mean Sq F value   Pr(>F)
temp         3  434.1  144.69  18.737 1.76e-06 ***
```

```
prep         2  128.4   64.19   8.313  0.00181 **
temp:prep    6   75.2   12.53   1.622  0.18426
Residuals   24  185.3    7.72
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> model2 <- aov(strength~ temp*prep*day,data=tensdata)
> summary(model2)
              Df Sum Sq Mean Sq
temp           3  434.1  144.69
prep           2  128.4   64.19
day            2   77.6   38.78
temp:prep      6   75.2   12.53
temp:day       6   20.7    3.44
prep:day       4   36.3    9.07
temp:prep:day 12   50.8    4.24
>
> model3 <- aov(strength~
prep+day+prep:day+temp+prep:temp+Error(day/prep),data=tensdata)
> summary(model3)

Error: day
    Df Sum Sq Mean Sq
day  2  77.56   38.78

Error: day:prep
         Df Sum Sq Mean Sq
prep      2 128.39   64.19
prep:day  4  36.28    9.07

Error: Within
          Df Sum Sq Mean Sq F value   Pr(>F)
temp       3  434.1  144.69  36.427 7.45e-08 ***
prep:temp  6   75.2   12.53   3.154   0.0271 *
Residuals 18   71.5    3.97
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> model.tables(model3,type="mean")
Tables of means
Grand mean

36.02778

 prep
prep
    1     2     3
35.67 38.50 33.92

 day
day
    1     2     3
34.75 35.25 38.08

 temp
temp
    1     2     3     4
31.22 34.56 37.89 40.44

 prep:day
    day
prep 1     2     3
   1 34.50 35.25 37.25
   2 38.75 37.25 39.50
   3 31.00 33.25 37.50

 prep:temp
    temp
prep 1     2     3     4
   1 29.67 34.67 39.33 39.00
   2 33.33 39.00 39.67 42.00
   3 30.67 30.00 34.67 40.33
```

a) What is the name of the design used in this experiment?

b) Is there a significant effect due to: (1) temperature; (2) preparation method at the 5% significance level?