4 Analysis of Variance and Design of Experiments

Preliminary Remark Analysis of variance (ANOVA) and design of experiments are both topics that are usually covered in separate lectures of about 30 hours. Here, we can only give a very brief overview. However, for many of you it may be worthwhile to study these topics in more detail later.

Analysis of variance addresses models where the response variable Y is a function of **categorical predictor variables** (so called **factors**). We have already seen how such predictors can be applied in a linear regression model. This means that analysis of variance can be viewed as a special case of regression modeling. However, it is worthwhile to study this special case separately. Analysis of variance and linear regression can be summarized under the term **linear model**.

Regarding design of experiments we only cover one topic, the **optimization of a response variable**. If time permits, we will also discuss some more general aspects.

4.1 Multiple Groups, One-Way ANOVA

a We observe g groups of values

$$Y_{hi} = \mu_h + E_{hi} \qquad i = 1, 2, \ldots, n_h; \qquad h = 1, 2, \ldots, g,$$

where $E_{hi} \sim \mathcal{N}(0, \sigma^2)$, independent.

The question of interest is whether there is a difference between the μ_h's.

b **Null hypothesis** $H_0 : \mu_1 = \mu_2 = \ldots = \mu_g$.

Alternative $H_A : \mu_h \neq \mu_k$ for at least one pair (h, k).

Test statistic
Based on the average of each group $\overline{Y}_{h.} = \frac{1}{n_h} \sum_{i=1}^{n_h} Y_{hi}$ we get the "mean squared error between the different groups"

$$\text{MSG} = \frac{1}{g-1} \sum_{h=1}^{g} n_h (\overline{Y}_{h.} - \overline{Y}_{..})^2.$$

This can be compared to the "mean squared error within the groups"

$$\text{MSE} = \frac{1}{n-g} \sum_{h,i} (Y_{hi} - \overline{Y}_{h.})^2,$$

leading to the test statistics of the **F-test**:

$$T = \frac{MSG}{MSE},$$

which follows an F-distribution with $g-1$ and $n-g$ degrees of freedom under H_0.

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
Treatment	4	520.69	130.173	1.508	0.208
Error	77	6645.15	86.301		
Total	81	7165.84			

Table 4.1.b: Example of an ANOVA table.

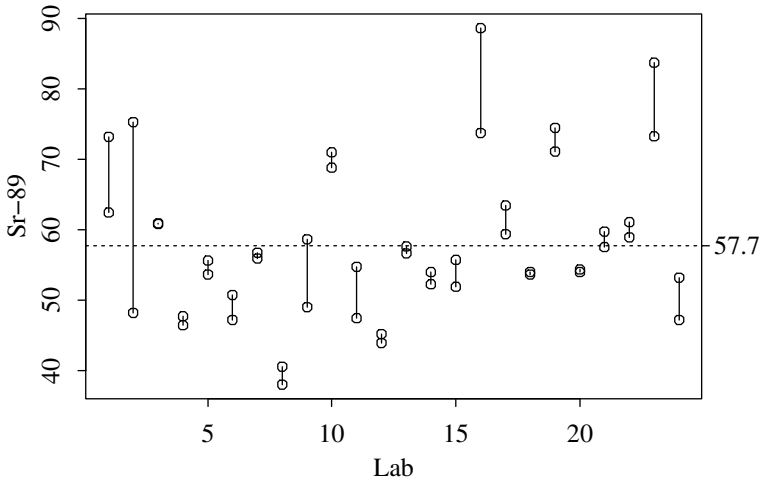

Figure 4.2.a: Sr-89-values for the 24 laboratories

c **Non-parametric tests** (if errors are not normally distributed):
"Kruskal-Wallis-Test", based on the ranks of the data.
For $g = 2$ groups: "Wilcoxon-Mann-Whitney-Test", also called "U-Test".

4.2 Random Effects, Ring Trials

Example **a** **Strontium in Milk** Figure 4.2.a illustrates the results of a ring trial (an inter-laboratory comparison) to determine the concentration of the radioactive isotope **Sr-89 in milk** (the question was of great interest after the Chernobyl accident). In 24 laboratories in Germany two runs to determine this quantity in artificially contaminated milk were performed. For this special situation the "true value" is known: it is 57.7 Bq/l. Source: G. Haase, D. Tait und A. Wiechen: "Ergebnisse der Ringanalyse zur Sr-89/Sr-90-Bestimmung in Milch im Jahr 1991". Kieler Milchwirtschaftliche Forschungsberichte 43, 1991, S. 53-62).

Figure 4.2.a shows that the two measurements of the same laboratory are in general much more similar than measurements between different laboratories.

b **Model**: $Y_{hi} = \mu + A_h + E_{hi}$. A_h **random**, $A_h \sim \mathcal{N}(0, \sigma_A^2)$.

Special quantities can tell us now how far two measurements can be from each other such that it is still safe to assume that the difference is only random.

- Comparisons **within** laboratory: **"Repeatability"** $2\sqrt{2} \cdot \widehat{\sigma}_E$,
- Comparisons **between** laboratories: **"Comparability"** $2\sqrt{2(\widehat{\sigma}_E^2 + \widehat{\sigma}_A^2)}$

4.3 Two and More Factors

Example a **Fisher's Potato Crop Data** Sir Ronald A. **Fisher** who established ANOVA (and many other things), used to work in the agricultural research center in Rothamstead, England. In an experiment to increase the yield of potatoes, the influence of **two treatment factors**, the addition of ammonium- and potassium-sulphate (each having 4 levels: 1, 2, 3, 4), was studied. Figure 4.3.a illustrates the data. Source: T. Eden and R. A. Fisher, Studies in Crop Variation. VI. Experiments on the Response of the Potato to Potash and Nitrogen, J. Agricultural Science, 19, 201-213, 1929; available through Bennett, 1971.

		Potassium-sulphate K			
		1	2	3	4
Ammonium-sulphate N	1	404	308	356	439
	2	318	434	402	422
	3	456	544	484	504
	4	468	534	500	562

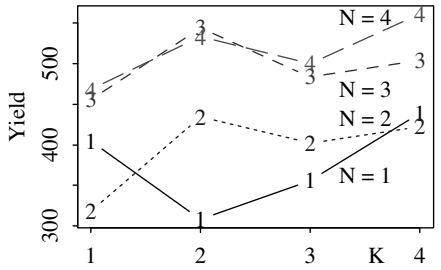

Figure 4.3.a: Fisher's Potato Crop Data.

b Model:

$$Y_{h,k} = \mu + \alpha_h + \beta_k + E_{hk} \ , \qquad \sum_h \alpha_h = 0 \quad \text{und} \quad \sum_k \beta_k = 0$$

c Estimates:

$$\widehat{\mu} = \overline{Y}_{..} \ , \qquad \widehat{\alpha}_h = \overline{Y}_{h.} - \overline{Y}_{..} \ , \qquad \widehat{\beta}_k = \overline{Y}_{.k} - \overline{Y}_{..} \ .$$

d **Tests**. Null-hypotheses: No influence of factor A (B). F-Tests. See table.

	DF	SS	MS	F	Pr(F)
N	3	59793	19931	10.84	0.0024
K	3	10579	3526	1.92	0.1973
Resid.	9	16552	1839		
Total	15	86924			

Table 4.3.d: ANOVA table for Fisher's potato crop data.

e **Interaction Effects** Model 4.3.b assumes that the effect of factor B is given by β_k, independent of the value of factor A. Or in other words, the model postulates that the effects of the two factors are additive. In general, so called **interaction effects** can occur. E.g., for fertilizers, further increasing one fertilizer is of little effect if another substance (fertilizer) is missing.

The general model for two factors with interaction effect can be written as

$$Y_{h,k} = \mu_{h,k} + E_{h,k} = \mu + \alpha_h + \beta_k + \gamma_{hk} + E_{h,k} \ .$$

Side constraints for the the interaction effect γ_{hk} are needed in order to obtain an identifiable model: $\sum_h \gamma_{hk} = 0$ for all k and $\sum_k \gamma_{hk} = 0$ for all h.

However, parameters can only be estimated if there are two or more observations for each combination of (h, k) (replicates).

f It's not difficult to extend model 4.3.b for **more than two factors**. The general model then also contains "interactions of higher order".

g For product development it's often necessary to check the effect of several (many) factors. In order to avoid too many experiments, it's often useful to restrict each factor to two levels and to avoid replicates. Such a series of experiments for k factors is called 2^k**-design** and will be discussed in more detail in the next section.

4.4 Response Surface Methods

Example a **Antibody Production** Large amounts of antibodies are obtained in biotechnological processes: Host animals (e.g. mice) are injected with modified cells that can produce the corresponding antibody. After a certain time these cells start to produce antibodies that can be collected in excreted fluid for further processing.

The cells can only produce antibodies if the immune system of the host animal is being weakened at the same time. This can be done with 4 factors. Moreover, it is believed that the amount of injected cells and their development stage has an influence on antibody production.

As there are no theoretical models for such complex biological processes, the relevant process factors have to be determined by an experiment. Such an experiment needs many mice, is time-intensive and usually costs a lot of money. Using a clever design, we can find out the important process factors with the lowest possible effort. That's where **statistical design of experiments** comes into play.

Two relevant process factors were identified in this study: the dose of Co^{60} gamma rays and the number of days between radiation and the injection of a pure oil. Now, the question is to find the levels for these two factors such that an **optimal amount** of antibodies is being produced by the modified cells.

b We have already seen a model which models a response variable Y that depends on two factors. It was

$$Y_{h,k} = \mu_{h,k} + E_{h,k} = \mu + \alpha_h + \beta_k + \gamma_{hk} + E_{h,k}\ , \qquad h,\, k = 1,\, 2.$$

If the two factors are based on continuous variables $x^{(1)}$, $x^{(2)}$, as is the case here with radiation dose and the number of days between radiation and injection, we have the corresponding general model

$$Y_i = h\left(x_i^{(1)}, x_i^{(2)}\right) + E_i,$$

(analogous for more than two factors). The function $h\left(x^{(1)}, x^{(2)}\right)$, which depends on $x^{(1)}$ and $x^{(2)}$, is the so-called **response surface**. Usually a quadratic polynomial (see below) in the variables $x^{(1)}$ and $x^{(2)}$ is used for h (sometimes the function h is available from theory). Once we have h, we can find the optimal setting $[x_0^{(1)}, x_0^{(2)}]$ of the process factors. Usually, h must be estimated from data.

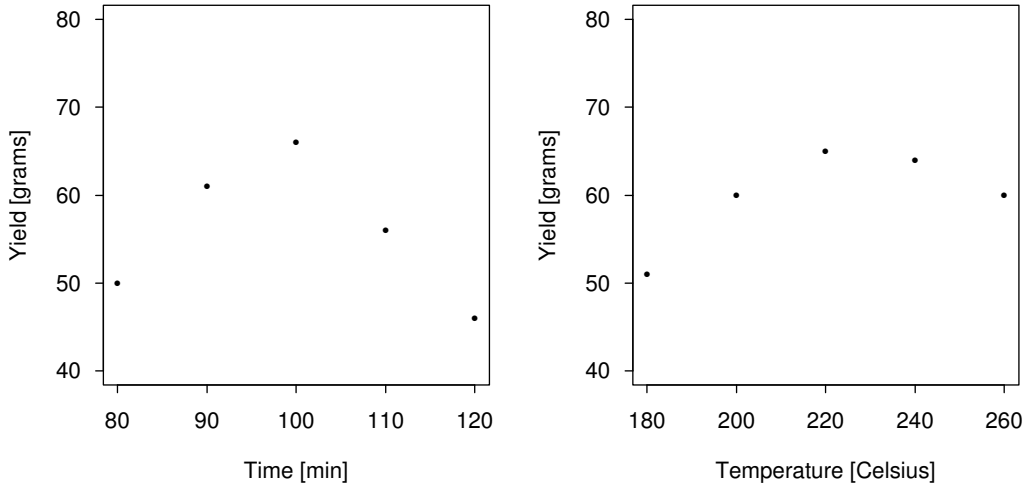

Figure 4.4.c: Varying the variables one by one. Left: Yield vs. reaction time, reaction temperature held constant at $220°C$. Right: Yield vs. reaction temperature, reaction time held constant at 100 minutes.

c A naive approach to find the optimum would be to optimize the variables **one by one**. The weakness of such an approach is now being illustrated with an artificial example.

Example d **Reaction Analysis** A chemist wants to maximize the yield of a chemical reaction by varying reaction time and reaction temperature. First, he performs an experiment where he uses a constant reaction temperature of $T = 220°C$ and reaction times 80, 90, 100, 110, and 120 minutes. Results are illustrated in Figure 4.4.d. According to this data, the maximum is attained with a reaction time of about 100 minutes.

In a second stage, reaction time is held constant at its optimal value of $t = 100$ minutes. Reaction temperature is varied at 180, 200, 220, 240, and $260°C$. Now, the conclusion is that maximal yield is attained with a reaction temperature of about $220°C$. This is not too far away from the value that was used in the first stage. Hence, the final conclusion is that the maximal yield of about 65 grams is attained using a reaction time of about 100 minutes and a reaction temperature of about $220°C$.

e To see that this conclusion is wrong, we have to make use of a two-dimensional view. Let us put time on the x- and temperature on the y-axis. Yield is illustrated by the corresponding contours (Figure 4.4.e). In this example, maximal yield of about 70 grams is attained with a reaction time of about 85 minutes and a reaction temperature of about $270°C$.

The approach of "varying the variables one by one" is misleading because it tacitly assumes that the maximal value of one variable is independent of the other ones. This assumption is usually not fulfilled.

f Even though the original setting of the process variables was "far away" from the optimal value, an appropriate sequence of well chosen experimental set-ups leads to the optimum. For that purpose, we start with a so called **first-order design**, a 2^k-design with additional measurements in the center. Experience from earlier experiments should guide us in selecting appropriate levels for the factors.

Example g **Reaction Analysis (cont'd)** From earlier experiments we know that a reaction temperature of $140°C$ and a reaction time of 60 minutes gives good results. Now we want to vary reaction time by 10 minutes and reaction temperature by $20°C$. The corre-

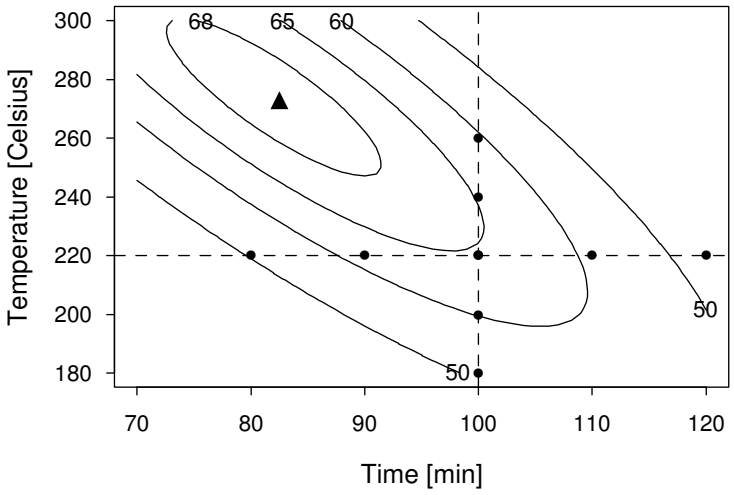

Figure 4.4.e: A hypothetical response surface, illustrated by contours in the diagram reaction temperature vs. reaction time.

sponding first-order design and the corresponding measurement results can be found in Table 4.4.g. Usually, coded variables are used in literature. They can also be found in Table 4.4.g.

	Variable in original units		Variable in coded units		Yield
Run	Temperature [°C]	Time [min]	Temperature	Time	Y [grams]
1	120	50	–1	–1	52
2	160	50	+1	–1	62
3	120	70	–1	+1	60
4	160	70	+1	+1	70
5	140	60	0	0	63
6	140	60	0	0	65

Table 4.4.g: First-order design and measurement results for the example "Reaction Analysis". The single experiments (runs) were performed in random order : 5, 4, 2, 6, 1, 7, 3.

h Because the response surface h (see 4.4.b) is unknown, we approximate it with the simplest possible surface, a plane. Hence, we have the model

$$Y_i = \theta_0 + \theta_1 x_i^{(1)} + \theta_2 x_i^{(2)} + E_i,$$

which has to be fitted to the data. We have already seen how the parameter estimates can be obtained.

The fitted plane, the so called **first-order response surface**, is given by

$$\widehat{y} = \widehat{\theta}_0 + \widehat{\theta}_1 x^{(1)} + \widehat{\theta}_2 x^{(2)}.$$

Of course, this is only an approximation of the real response surface.

Example i **Reaction Analysis (cont'd)** The parameters with respect to the coded variables can be found in Table 4.4.i.

```
             Estimate  Std. Error  t value     Pr(> |t|)
(Intercept)    62.000       0.882    70.30  6.3e-06 ***
xt1             5.000       1.080     4.63      0.019 *
xt2             4.000       1.080     3.70      0.034 *
```

Table 4.4.i: Estimated coefficients for the coded variables in the example "Reaction Analysis".

j On the first-order response surface we can find those points $[x^{(1)},\ x^{(2)}]^T$ which have a constant yield $\widehat{y} = \widehat{y}_0$. From equation

$$\widehat{y}_0 = \widehat{\theta}_0 + \widehat{\theta}_1 x^{(1)} + \widehat{\theta}_2 x^{(2)}$$

we find the straight line

$$x^{(2)} = \frac{\widehat{y}_0 - \widehat{\theta}_0 - \widehat{\theta}_1 x^{(1)}}{\widehat{\theta}_2}$$

with slope $b = -\widehat{\theta}_1/\widehat{\theta}_2$ and intercept $a = (\widehat{y}_0 - \widehat{\theta}_0)/\widehat{\theta}_2$. Orthogonal to this straight line is the **direction of steepest ascent (descent)**. This straight line has slope $\widehat{\theta}_2/\widehat{\theta}_1$. The two-dimensional vector $[\widehat{\theta}_1, \widehat{\theta}_2]^T$ is called estimated **gradient**; this direction is the fastest way to get large values of $\widehat{y}$.

Of course we also get large values when following any direction that is "close" to the gradient.

k Observations that are in the center of the 2^k-design have no influence on the estimates of the parameters $\theta_1, \ldots, \theta_k$ and hence no influence on the estimated gradient, either. This can be seen from the normal equations.

But why should we do experiments in the center?

- It's possible to estimate the measurement error without using the assumption that the plane is a good approximation of the true response surface if several observations are available in the center.
- Possible curvature of the true response surface can be detected. If there is no curvature and if the plane is a "good" approximation of the true response surface in the range of the experimental set-up, the average of the observations in the center, $\overline{Y}_c$, and the average of the observations of the 2^k-design, $\overline{Y}_f$, are estimates of the mean of Y for the set-up in the center. Hence, they should be "more or less equal". If the difference is obviously different from zero it's a hint that there is curvature.

A statistical test for curvature is as follows: The empirical variance s^2 that was estimated from the n_c observations in the center can be used to determine the standard deviation of the difference. The variance of $\overline{Y}_c$ can be estimated by s^2/n_c and the variance of $\overline{Y}_f$ by $s^2/2^k$. Because $\overline{Y}_c$ and $\overline{Y}_f$ are independent, the variance of the difference $\overline{Y}_c - \overline{Y}_f$ is estimated by $s^2(1/n_c + 1/2^k)$. Now we can perform a t-test or we can construct the corresponding confidence interval: If the interval

$$\overline{Y}_c - \overline{Y}_f \pm q_{0.975}^{t_{n_c-1}} \cdot \sqrt{s^2(1/n_c + 1/2^k)}$$

does not cover zero, the difference is statistically different from zero.

If we face relevant curvature of the true response surface (in the range of the experimental set-up), a linear approximation is not appropriate. Hence, we may also have problems determining the direction of steepest ascent. Usually we will use a second-order response surface (see below) for such situations.

l If the measurements of the center observations are all performed in a row, we face the danger that the observed variation (measured by s^2) is not really the variation between "independent" observations. Usually we will get significant curvature, even for cases where the response surface is a plane.

In general, it's **important to randomize the different experimental set-ups** – even though this usually needs much more effort because we always have to arrange a new setting for each new run.

m If it's plausible to use a linear response surface, we will search for an optimum along the direction of steepest ascent. Along

$$\begin{bmatrix} x_0^{(1)} \\ x_0^{(2)} \end{bmatrix} + k \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix}$$

we will perform additional experiments for $k = 1, 2, \ldots$ until yield starts to decrease. $[x_0^{(1)},\ x_0^{(2)}]^T$ is the point in the center of our experimental design and $[c^{(1)},\ c^{(2)}]^T$ is the direction of steepest ascent.

Example n **Reaction Analysis (cont'd)** The first-order response surface is

$$\widehat{y} = 62 + 5\widetilde{x}^{(1)} + 4\widetilde{x}^{(2)} = 3 + 0.25 \cdot x^{(1)} + 0.4 \cdot x^{(2)},$$

where $[\widetilde{x}^{(1)},\ \widetilde{x}^{(2)}]^T$ are the coded x-values (taking values ± 1) (see Table 4.4.i). Note that the gradient for the coded and the non-coded (original) x-values lead to different directions of steepest ascent. An other ascent direction can be identified by observing that individually increasing temperature by $4°C$ or time by 2.5 minutes both leads to a yield increase of 1 gram.

Further experiments are now performed along

$$\begin{bmatrix} 140 \\ 60 \end{bmatrix} + k \cdot \begin{bmatrix} 25 \\ 10 \end{bmatrix},$$

which corresponds to the steepest ascent direction with respect to the coded x-values, (see Table 4.4.n).

	Temperature $[°C]$	Time [min]	Y
1	165	70	72
2	190	80	77
3	215	90	79
4	240	100	76
5	265	110	70

Table 4.4.n: Experimental design and measurement results for experiments along the steepest ascent direction for the example "Reaction Analysis".

Based on the results in Table 4.4.n (plot the profile of yield vs. runs), the optimum should be in the neighborhood of a reaction temperature of $215°C$ and a reaction time of 90 minutes.

o Once the optimum along the gradient is identified, a further first-order design experiment can be performed (around the optimum) to get a new gradient. However, in general the hope is that we are already close to the optimal solution and we will continue as illustrated in the next section.

4.5 Second-Order Response Surfaces

a Once we are close to the optimal solution, the estimated plane will be nearly parallel to the $(x^{(1)}, x^{(2)})$-plane. Hence, $\widehat{\theta}_1$ and $\widehat{\theta}_2$ will be nearly zero. We expect that the optimal solution is a (flat) peak in the range of our experimental set-up and hence we expect the difference $\overline{Y}_c - \overline{Y}_f$ to be significantly different from zero. Such a peak can be modelled by a second-order polynomial:

$$Y_i = \theta_0 + \theta_1 x_i^{(1)} + \theta_2 x_i^{(2)} + \theta_{11}(x_i^{(1)})^2 + \theta_{22}(x_i^{(2)})^2 + \theta_{12} x_i^{(1)}\, x_i^{(2)} + E_i$$

b However, the 2^k-design does not contain enough data to estimate the parameters of the second-order polynomial. The reason is that we need at least 3 levels for each factor. There are now several ways of expanding our original design. The more levels we have for each factor, the better we can estimate the curvature. So called **rotatable central composite designs** (also called second-order central composite designs) are very famous. As can be seen from the graphical representation in Figure 4.5.b we can get such a design by extending our original first-order design. In total we have 9 different experimental set-ups if we have two predictor variables. All points (except the center point) have the **same distance from the center** $(0, 0)$. Five levels are used for each factor. If we use replicates at $(0, 0)$ we get a more precise estimate of the quadratic part of the model.

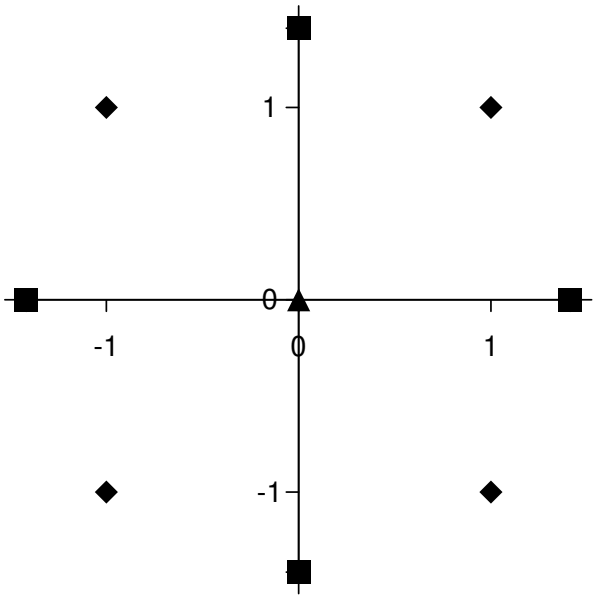

Figure 4.5.b: Rotatable central composite design. It consists of a 2^2-design (◆) with additional experiments in the center and along the axes (■).

Example c **Reaction Analysis (cont'd)** A rotatable central composite design was applied. The results can be found in Table 4.5.c.

Run	Variables in original units Temperature [$°C$]	Time [min]	Variables in coded units Temperature	Time	Yield Y [grams]
1	195	80	-1	-1	78
2	235	80	$+1$	-1	76
3	195	100	-1	$+1$	72
4	235	100	$+1$	$+1$	75
5	187	90	$-\sqrt{2}$	0	74
6	243	90	$+\sqrt{2}$	0	76
7	215	76	0	$-\sqrt{2}$	77
8	215	104	0	$+\sqrt{2}$	72
9	215	90	0	0	80

Table 4.5.c: Rotatable central composite design and measurement results for the example "Reaction Analysis".

The parameter estimates lead to the estimated **second-order response surface**:

$$\begin{aligned}\widehat{y} &= \widehat{\theta}_0 + \widehat{\theta}_1 x^{(1)} + \widehat{\theta}_2 x^{(2)} + \widehat{\theta}_{11}(x^{(1)})^2 + \widehat{\theta}_{22}(x^{(2)})^2 + \widehat{\theta}_{12} x^{(1)}\, x^{(2)} \\ &= -278 + 2.0 \cdot x^{(1)} + 3.2 \cdot x^{(2)} + 0.0060 \cdot (x^{(1)})^2 + 0.026 \cdot (x^{(2)})^2 + 0.006 \cdot x^{(1)}\, x^{(2)}.\end{aligned}$$

d Depending on the parameters, a second-order response surface can take different shapes. The most important ones are those that have a maximum (minimum) or a saddle (rather rare). A schematic contour plot of these two types is illustrated in Figure 4.5.d.

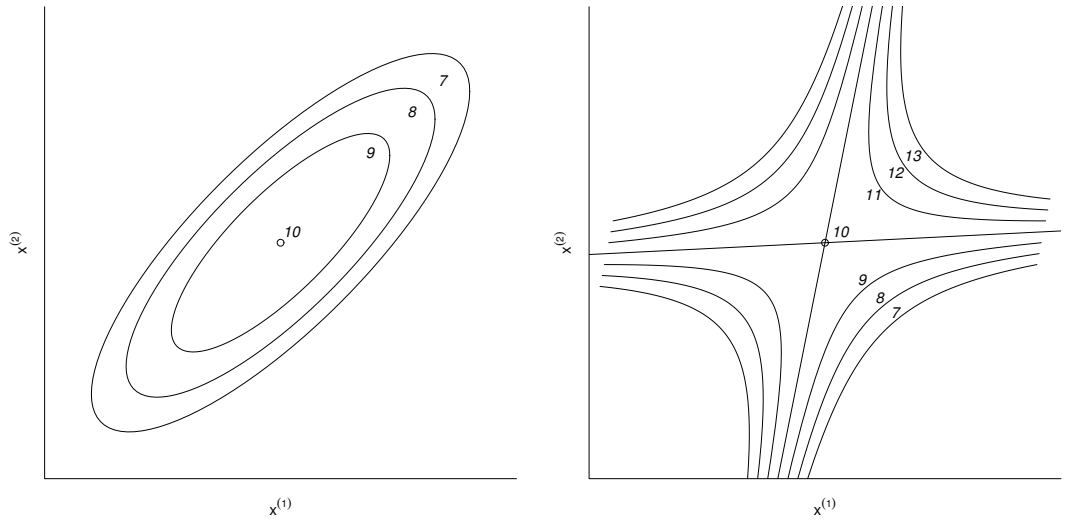

Figure 4.5.d: Contour plots of second-order response surfaces with a maximum (left) and a saddle (right).

Surfaces with a maximum (minimum) don't need further explanations: Once we leave the optimum in any direction, yield Y is decreasing (increasing). For a saddle, it depends on the direction whether yield Y increases or decreases. Hence, the surface is like a horse saddle.

e It's possible to find an analytical solution for the critical point by calculating partial derivatives. In the critical point they have to be zero:

$$\begin{aligned} \frac{\partial \widehat{Y}}{\partial x^{(1)}} &= \widehat{\theta}_1 + 2\,\widehat{\theta}_{11}\,x_0^{(1)} + \widehat{\theta}_{12}\,x_0^{(2)} = 0 \\ \frac{\partial \widehat{Y}}{\partial x^{(2)}} &= \widehat{\theta}_2 + 2\,\widehat{\theta}_{22}\,x_0^{(2)} + \widehat{\theta}_{12}\,x_0^{(1)} = 0\,. \end{aligned}$$

Solving this linear equation system for $x_0^{(1)}$ and $x_0^{(2)}$ leads us to

$$x_0^{(1)} = \frac{\widehat{\theta}_{12}\,\widehat{\theta}_2 - 2\,\widehat{\theta}_{22}\,\widehat{\theta}_1}{4\,\widehat{\theta}_{11}\,\widehat{\theta}_{22} - \widehat{\theta}_{12}^2} \qquad x_0^{(2)} = \frac{\widehat{\theta}_{12}\,\widehat{\theta}_1 - 2\,\widehat{\theta}_{11}\,\widehat{\theta}_2}{4\,\widehat{\theta}_{11}\,\widehat{\theta}_{22} - \widehat{\theta}_{12}^2}\,.$$

Example f **Reaction Analysis (cont'd)** We can directly read the critical values off the contour plot in Figure 4.5.f: ($220°C$, 85 minutes).

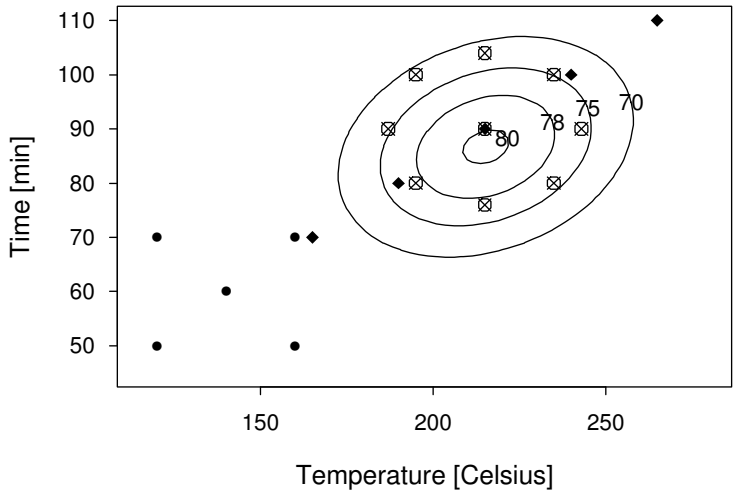

Figure 4.5.f: Estimated response surface and experimental designs for the example "Reaction Analysis". The first-order design is marked with ●, the experiments along the steepest ascent with ◆ and those of the rotatable central composite design with ⊠.

Example g **Antibody production (cont'd)** Let's come back to our original example. A radioactive dose of 200 rads and a time of 14 days is used as starting point. Around this point we use a first-order design, see Table 4.5.g.

Based on the measurements in the center we can calculate the standard deviation of the error term: $\widehat{\sigma} = 53.9$. Now we check whether there is significant curvature. The confidence interval for the difference can be calculated as outlined in 4.4.k:

$$\begin{aligned} \overline{Y}_c - \overline{Y}_f \pm q_{0.975}^{t_{n_c-1}} \cdot \sqrt{s^2(1/n_c + 1/2^k)} &= 589 - 335 \pm 4.30 \cdot 41.1 \\ &= [77,\ 431]. \end{aligned}$$

As this interval does not cover 0, the difference is statistically different from zero on the 5% level. As yield decreases at the border of our experimental set-up, we conjecture that the optimal value must lie somewhere within our set-up range.

Run	Variables in original units RadDos [rads]	Time [days]	Variables in coded units RadDos	Time	Yield Y
1	100	7	-1	-1	207
2	100	21	-1	$+1$	257
3	300	7	$+1$	-1	306
4	300	21	$+1$	$+1$	570
5	200	14	0	0	630
6	200	14	0	0	528
7	200	14	0	0	609

Table 4.5.g: First-order design and measurement results for the example "Antibody Production".

Hence, we expand our design to a rotatable central composite design by doing additional measurements (see Table 4.5.g).

Run	Variables in original units RadDos [rads]	Time [days]	Variables in coded units RadDos	Time	Yield Y
8	200	4	0	$-\sqrt{2}$	315
9	200	24	0	$+\sqrt{2}$	154
10	59	14	$-\sqrt{2}$	0	100
11	341	14	$+\sqrt{2}$	0	513

Table 4.5.g: Rotatable central composite design and measurement results for the example "Antibody Production".

The estimated response surface is

$$\begin{aligned}\widehat{Y} &= -608.4 + 5.237 \cdot \text{RadDos} + 77.0 \cdot \text{Time} \\ &\quad -0.0127 \cdot \text{RadDos}^2 - 3.243 \cdot \text{Time}^2 + 0.0764 \cdot \text{RadDos} \cdot \text{Time}.\end{aligned}$$

We can identify the optimal conditions in the contour plot (Figure 4.5.g):
$\text{RadDos}_{opt} \approx 250$ rads and $\text{time}_{opt} \approx 15$ days.

h **Summary** To find the optimal setting of our variables (leading to maximal yield) we have to iteratively do experiments using special designs.

- If we are still "far away" from the optimum, we use first-order designs and we estimate the corresponding first-order response surface (a plane).
- On the estimated response surface we determine an ascent direction. Along that direction we do additional experiments until the response variable decreases again ("extrapolation").
- Further first-order experiments may be performed (hence we repeat the last two steps).
- As soon as we are close to the optimum, we perform (e.g.) a rotatable central composite design (or we expand our first-order design) to estimate the second-order response surface. The optimum on that response surface can be determined either analytically or graphically.

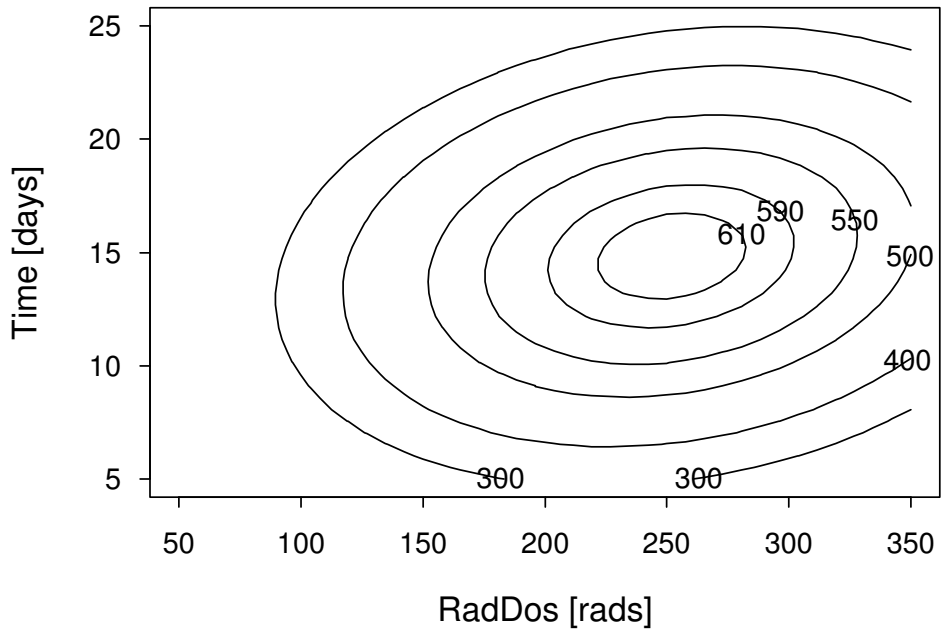

Figure 4.5.g: Estimated second-order response surface for the example "antibody production".

4.6 Experimental Designs, Robust Designs

a Here we discussed two types of experimental designs in more detail. Of course there are books full of other designs that are useful for various scopes. The subject is called **design of experiments**.

It may be worthwhile to note that for situations where little is known about the influence of the predictors on the response, so called "**screening designs**" can be very useful. They allow a (rough) analysis of k factors with less than 2^k experiments.

b The idea of "**robust product design**" is that products should have constant quality even if production conditions vary. To reach this goal we do not only have to optimize the expected quality (or yield or other response variables) but also the variability. There are special designs for that purpose, e.g. the **Taguchi designs**.

4.7 Further Reading

- **ANOVA and Design of Experiments** Short overviews of simple ANOVA models can be found in Hartung, Elpelt and Klösener (2002, Chap. XI) and Sachs (2004, Chap. 7). Linder and Berchtold (1982) give a more detailed introduction.
- Applied books about ANOVA and design of experiments are the famous book of Box, Hunter and Hunter (1978) and the book of Daniel (1976).
- A special book that uses unusual ways to introduce known (and unknown) methods with focus an explorative analysis is Hoaglin, Mosteller and Tukey (1991).
- A classical mathematically oriented book about ANOVA is Scheffe (1959).
- **Design of Experiments** Federer (1972, 1991) is an introduction to statistics where design of experiments often takes center stage. More details can be found in the already mentioned book of Box et al. (1978) but also in Mead (1988).
- A systematic overview of experimental design can be found in Petersen (1985).
- A few books discuss topics about practical application of statistics that can't

be dealt with mathematics. Recommendations are Boen and Zahn (1982) and Chatfield (1996).

- **Response Surfaces** An introduction to response surfaces is available in Box et al. (1978, Chapter 15) or in Hogg and Ledolter (1992).

  Box and Draper (1987) and Myers and Montgomery (1995) cover more details.