# SAS-Exercise on GLM

The study has observations of the reflection from a hyper spectral laser taken on two pieces of lard (*spæk* in danish). On each piece of lard two repeated measurements at three different positions were taken. The three positions were picked at random. The observations were each done at five different wavelengths in the range from 500 to 910 nm. Based on an image taken at each wavelength a slope feature is extracted. The slope feature is the slope of the log-log curve of the intensity profile through the image. (Data: By courtesy from Camilla Himmelstrup. Details on the measurements may be found in: Depth Analysis of Food Structures: Hyperspectral Subsurface Laser Scattering by Otto Nielsen et al.). Below the measurement set up is indicated.

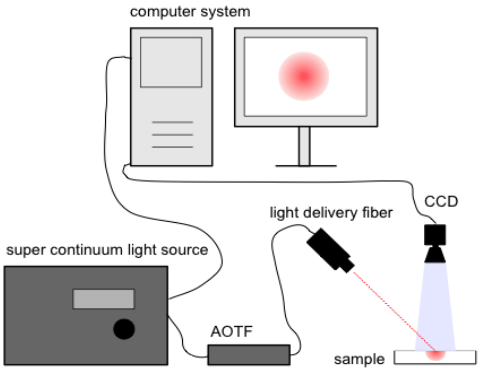


Figure 2: Experimental set up.

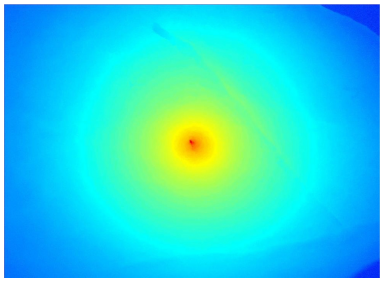

Figure 1: Image example. To enhance the differences in the image, it is the logarithm of the image that is shown. Through the center (the incident point of the laser) the intensity profile is found.

**Before getting started:**

The two data files can be accessed through Campusnet. It is important that you save the files as '.sas7bdat'-files in the stat2data-folder that you created in the first lecture of the semester.

**Data:**

The data is stored in two datasets, STAT2.larduni and STAT2.lardmul, and a printout of the two datasets is shown in the sequel. It follows that in 'larduni' the data is represented as a sequence of 60 univariate measurements, in 'lardmul' as a sequence of 12 5-dimensional variables each representing all measurements taken at a given position and replication number. We assume that the measurements in 'larduni' may be considered as realizations of independent random variables with the same variance. This assumption is not necessarily reasonable, since we would expect that there might be a correlation between repeated measurements taken at the same position. But in the analyses we will perform, it is assumed that such a correlation is zero. We shall now perform several statistical tests in order to investigate whether and how the measurements vary over the two pieces of lard.

# The data in STAT2.larduni

| Obs | Piece | posi | rep | Wave-length | Slope |
|---|---|---|---|---|---|
| **1** | 1 | 1 | 1 | 500 | -8.3873 |
| **2** | 1 | 1 | 2 | 500 | -8.3255 |
| **3** | 1 | 1 | 1 | 580 | -7.0316 |
| **4** | 1 | 1 | 2 | 580 | -7.0719 |
| **5** | 1 | 1 | 1 | 750 | -4.3492 |
| **6** | 1 | 1 | 2 | 750 | -4.2761 |
| **7** | 1 | 1 | 1 | 830 | -4.2591 |
| **8** | 1 | 1 | 2 | 830 | -4.1818 |
| **9** | 1 | 1 | 1 | 910 | -4.4846 |
| **10** | 1 | 1 | 2 | 910 | -4.403 |
| **11** | 1 | 2 | 1 | 500 | -8.2176 |
| **12** | 1 | 2 | 2 | 500 | -8.2486 |
| **13** | 1 | 2 | 1 | 580 | -6.5513 |
| **14** | 1 | 2 | 2 | 580 | -6.6171 |
| **15** | 1 | 2 | 1 | 750 | -3.9763 |
| **16** | 1 | 2 | 2 | 750 | -3.9667 |
| **17** | 1 | 2 | 1 | 830 | -3.7754 |
| **18** | 1 | 2 | 2 | 830 | -3.7659 |
| **19** | 1 | 2 | 1 | 910 | -3.8567 |
| **20** | 1 | 2 | 2 | 910 | -3.8362 |
| **21** | 1 | 3 | 1 | 500 | -8.291 |
| **22** | 1 | 3 | 2 | 500 | -8.2584 |
| **23** | 1 | 3 | 1 | 580 | -6.5804 |
| **24** | 1 | 3 | 2 | 580 | -6.5742 |
| **25** | 1 | 3 | 1 | 750 | -3.8994 |
| **26** | 1 | 3 | 2 | 750 | -3.883 |
| **27** | 1 | 3 | 1 | 830 | -3.7068 |
| **28** | 1 | 3 | 2 | 830 | -3.7031 |
| **29** | 1 | 3 | 1 | 910 | -3.738 |
| **30** | 1 | 3 | 2 | 910 | -3.7199 |

| Obs | Piece | posi | rep | Wave-length | Slope |
|---|---|---|---|---|---|
| **31** | 2 | 1 | 1 | 500 | -6.9975 |
| **32** | 2 | 1 | 2 | 500 | -7.0188 |
| **33** | 2 | 1 | 1 | 580 | -6.3663 |
| **34** | 2 | 1 | 2 | 580 | -6.3503 |
| **35** | 2 | 1 | 1 | 750 | -4.3607 |
| **36** | 2 | 1 | 2 | 750 | -4.3366 |
| **37** | 2 | 1 | 1 | 830 | -4.3052 |
| **38** | 2 | 1 | 2 | 830 | -4.3008 |
| **39** | 2 | 1 | 1 | 910 | -4.9364 |
| **40** | 2 | 1 | 2 | 910 | -4.9185 |
| **41** | 2 | 2 | 1 | 500 | -7.4082 |
| **42** | 2 | 2 | 2 | 500 | -7.4082 |
| **43** | 2 | 2 | 1 | 580 | -6.59 |
| **44** | 2 | 2 | 2 | 580 | -6.59 |
| **45** | 2 | 2 | 1 | 750 | -4.5574 |
| **46** | 2 | 2 | 2 | 750 | -4.5574 |
| **47** | 2 | 2 | 1 | 830 | -4.6532 |
| **48** | 2 | 2 | 2 | 830 | -4.6532 |
| **49** | 2 | 2 | 1 | 910 | -5.3502 |
| **50** | 2 | 2 | 2 | 910 | -5.3502 |
| **51** | 2 | 3 | 1 | 500 | -7.4037 |
| **52** | 2 | 3 | 2 | 500 | -7.3194 |
| **53** | 2 | 3 | 1 | 580 | -6.6295 |
| **54** | 2 | 3 | 2 | 580 | -6.3319 |
| **55** | 2 | 3 | 1 | 750 | -4.5518 |
| **56** | 2 | 3 | 2 | 750 | -4.3202 |
| **57** | 2 | 3 | 1 | 830 | -4.6628 |
| **58** | 2 | 3 | 2 | 830 | -4.4499 |
| **59** | 2 | 3 | 1 | 910 | -5.3275 |
| **60** | 2 | 3 | 2 | 910 | -5.1351 |

**Tasks on `larduni`:**

Analyze models of the form

```
class piece posi  rep;
model slope=piece*posi wavelength(piece posi)
wavelength*wavelength(piece posi) /noint solution;
```

Try to find a simpler model that gives a satisfactory description of the data!

# The data in STAT2.lardmul

| Obs | Piece | posi | rep | slo500 | slo580 | slo750 | slo830 | slo910 |
|---|---|---|---|---|---|---|---|---|
| **1** | 1 | 1 | 1 | -8.3873 | -7.0316 | -4.3492 | -4.2591 | -4.4846 |
| **2** | 1 | 1 | 2 | -8.3255 | -7.0719 | -4.2761 | -4.1818 | -4.4030 |
| **3** | 1 | 2 | 1 | -8.2176 | -6.5513 | -3.9763 | -3.7754 | -3.8567 |
| **4** | 1 | 2 | 2 | -8.2486 | -6.6171 | -3.9667 | -3.7659 | -3.8362 |
| **5** | 1 | 3 | 1 | -8.2910 | -6.5804 | -3.8994 | -3.7068 | -3.7380 |
| **6** | 1 | 3 | 2 | -8.2584 | -6.5742 | -3.8830 | -3.7031 | -3.7199 |
| **7** | 2 | 1 | 1 | -6.9975 | -6.3663 | -4.3607 | -4.3052 | -4.9364 |
| **8** | 2 | 1 | 2 | -7.0188 | -6.3503 | -4.3366 | -4.3008 | -4.9185 |
| **9** | 2 | 2 | 1 | -7.4082 | -6.5900 | -4.5574 | -4.6532 | -5.3502 |
| **10** | 2 | 2 | 2 | -7.4082 | -6.5900 | -4.5574 | -4.6532 | -5.3502 |
| **11** | 2 | 3 | 1 | -7.4037 | -6.6295 | -4.5518 | -4.6628 | -5.3275 |
| **12** | 2 | 3 | 2 | -7.3194 | -6.3319 | -4.3202 | -4.4499 | -5.1351 |

**Task on `lardmul`:**
Analyze models of the form:

```
model slo500 slo580 slo750 slo830 slo910=piece posi(piece)
rep(piece posi);
model Prin1-Prin5=piece posi(piece) rep(posi piece);
```

Here `Prin1-Prin5` are the five principal components of `slo500-slo910`.

You may also try to run a stepwise discriminant analysis like

```
proc stepdisc data=multipc;
var slo500 slo580 slo750 slo830 slo910;
class piece;
run;
```

where 'multipc' is the dataset containing the principal component scores.

*kco fecit 2. Nov 2012*