Sweat gland outliers

Mari Myllymäki, Mikko Kuronen, Aila Särkkä

2023-09-19

Data

In this example we analyse the sweat gland pattern of subject 203, a control. The original data (video) have been collected by Dr. Kennedy's group at the University of Minnesota by using the dynamic sweat test they have presented (Provitera et al. 2009. https://doi.org/10.1212/WNL.0b013e3181c7da4b). The point pattern analysed here was extracted by Kuronen et al. (2021, https://doi.org/10.1002/sim.8891) and provided at the repository https://github.com/mikkoku/SweatPaper accompanying the article.

We obtained the sweat gland pattern of subject 203 as follows:

```
library(spatstat)
library(tidyverse)
```

```
path <- "https://github.com/mikkoku/SweatPaper/raw/master/data"
glands <- read.table(file.path(path, "glands.csv"), header=TRUE, sep=",")
meta <- read.table(file.path(path, "meta.csv"), header=TRUE, sep=",")

s <- 17.5/2592 # scaling from pixels to millimeters
glands <- glands |> mutate(x=s*x, y=s*y)
meta <- meta |> mutate(across(3:6, ~ .x*s))

s203 <- glands |> filter(subjectid==203)
m203 <- meta |> filter(subjectid==203)

W <- owin(c(m203$x0, m203$x1),
          c(m203$y0, m203$y1))
X <- ppp(x=s203$x, y=s203$y, window=W)
par(mar=c(0,0,0,0), mgp=c(0,0,0))
plot(X, main="")
```
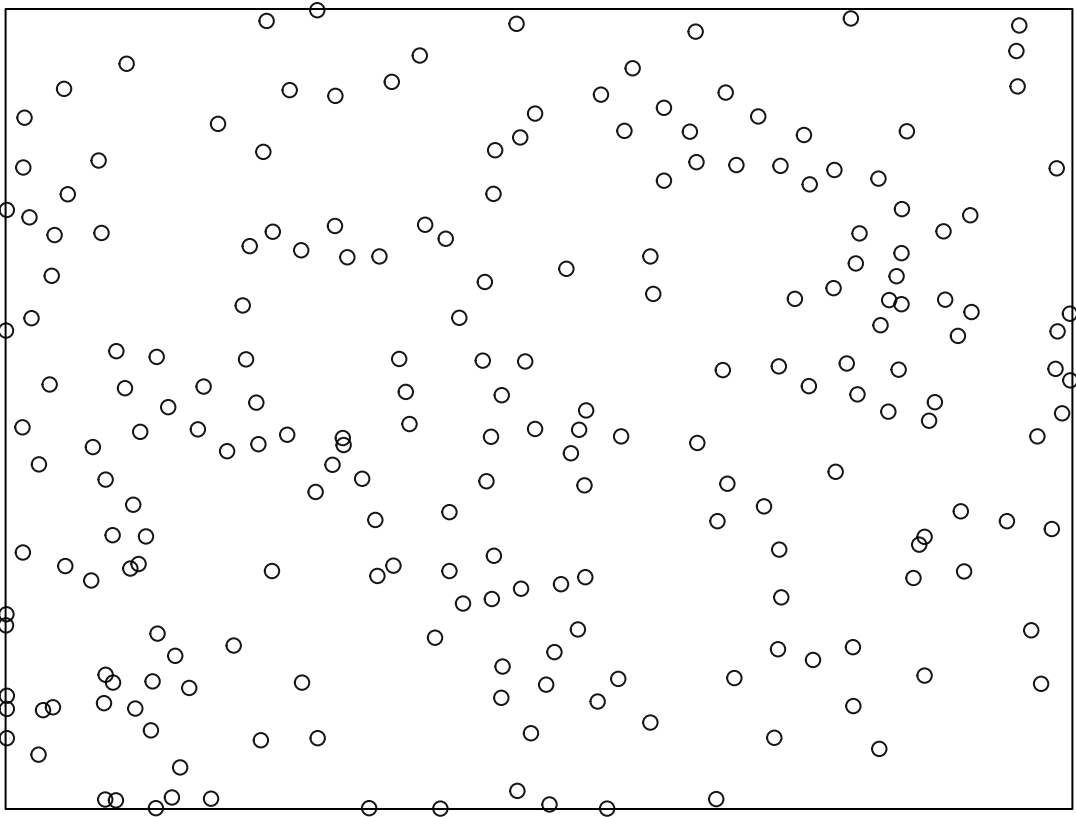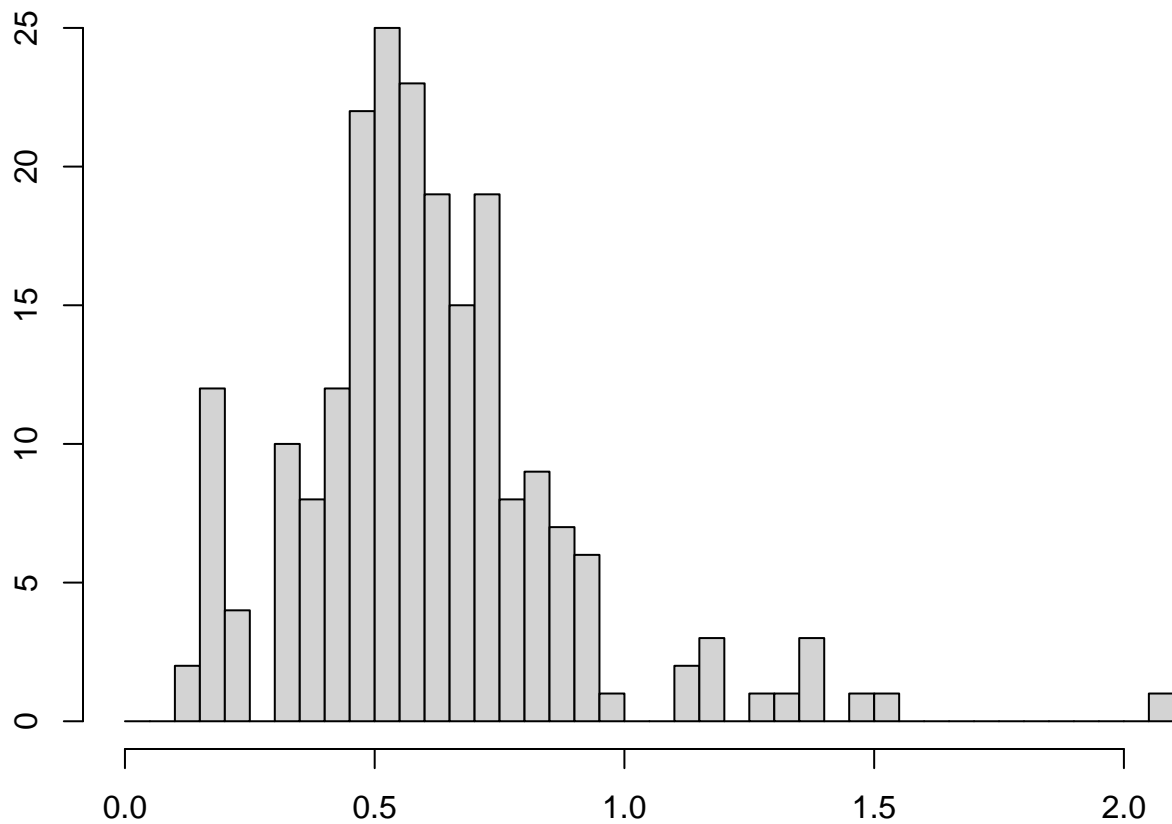
The nearest neighbor distance histogram shows some extraordinarily close pairs of points.

```r
par(mar=c(2.5,2.5,0.5,0.5))
hist(nndist(X), main="", breaks=seq(0, 2.1, by=0.05))
```

We plotted the point pattern and the points which had their nearest neighbor within distance 0.25 mm.

```
nnlimit <- 0.25
par(mfrow=c(1,1), mar=c(0,0,0,0), mgp=c(0,0,0))
plot(X, main="")
points(subset(X, nndist(X) <= nnlimit), col=2, pch=16)
points(subset(X, nndist(X) <= nnlimit), col=1, pch=1)
```
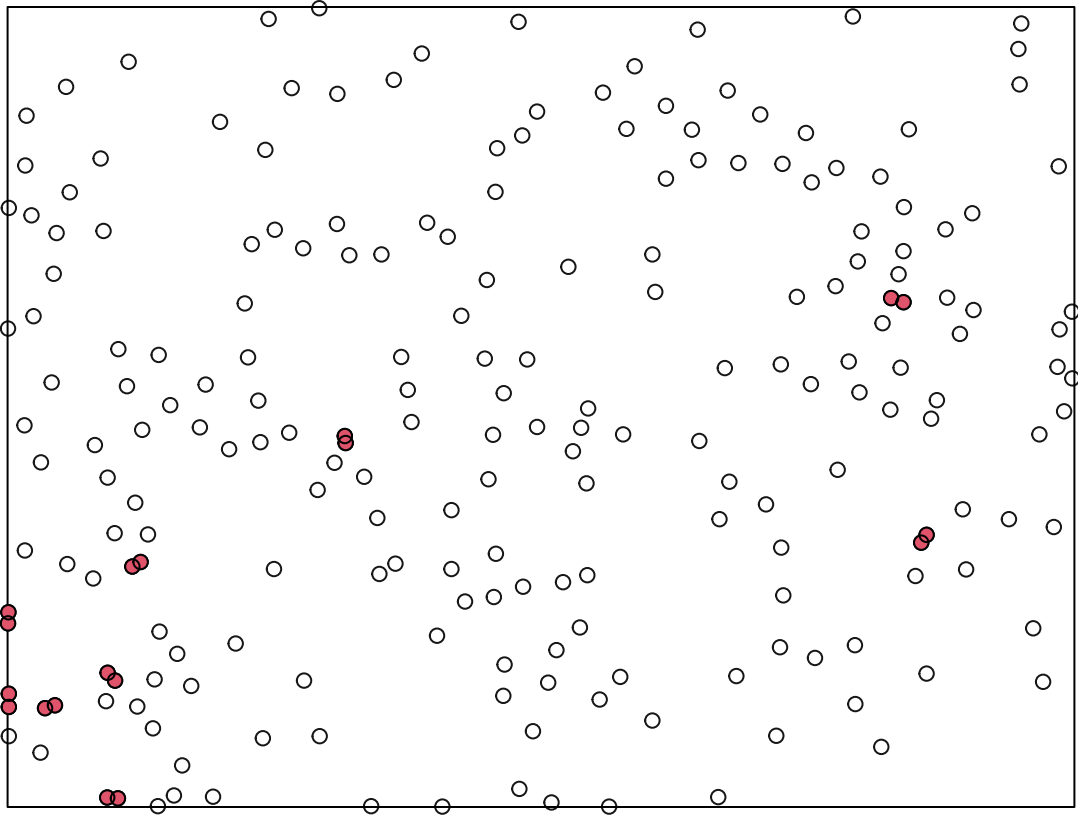
To see what the effect of the close pairs to the model fit is, we removed one of each pair of points that had their nearest neighbor within distance 0.25 mm.

```
Xr <- X
while(any(nndist(Xr) <= nnlimit)) {
  id <- which.min(nndist(Xr))
  Xr <- Xr[-id]
}
```

The pair correlation functions of the original and the filtered point patterns show difference in the short distances.

```
# pcf of the original pattern
X.pcf <- pcf(X, correction="translate", r=seq(0, 3, length=101))
# pcf of the filtered pattern
Xr.pcf <- pcf(Xr, correction="translate", r=seq(0, 3, length=101))

par(mar=c(2.5,2.5,0.5,0.5))
plot(X.pcf, legend=FALSE, main="", ylim=c(0,1.4))
plot(Xr.pcf, add=TRUE, col=2)
```
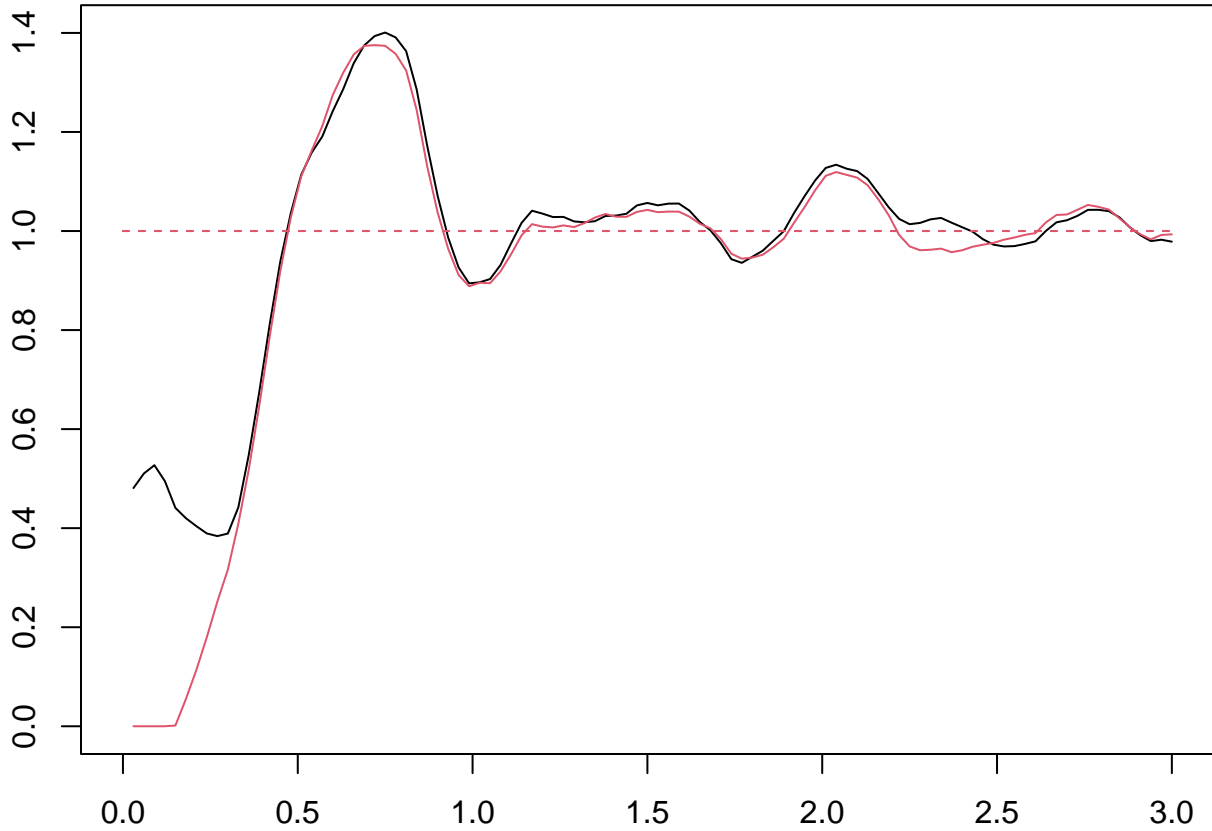
We then investigated how the outlier pairs affect the parameters a softcore Gibbs point process model which has the density function

$$f(\mathbf{x}) \propto \beta^{n(\mathbf{x})} \exp\left(-\sum_{i<j}\left(\frac{\sigma}{\|x_i - x_j\|}\right)^{2/\kappa}\right),$$

This model can be fitted by the R library `spatstat` (Baddeley et al. (2015)) using the `ppm` function. It is possible to obtain confidence intervals for the model parameters using the `summary` function, but only for internal parametrization. We obtained confidence intervals for β and σ through transformation.

```r
parameter_estimates <- function(M) {
  # Transform internal parametrization to user friendly parametrization
  c1 <- coef(summary(M, fine=TRUE))
  if(c1[2,3] < 0) c1[2,3] <- 0
  rbind(beta=exp(c1[1,c(1,3,4)]),
    sigma=M$interaction$par$sigma0*c1[2,c(1,3,4)]^(kappa/2))
}
kappa <- 0.2
M1 <- ppm(X~1, Softcore(kappa=kappa), correction="isotropic")
Mr <- ppm(Xr~1, Softcore(kappa=kappa), correction="isotropic")

M1

## Stationary Soft core process
## Fitted to point pattern dataset 'X'
##
## First order term:  beta = 0.9531694
##
```

```
## Exponent kappa:  0.2
## Initial approximation to sigma:  0.1155678
## Fitted interaction parameter sigma:  0.103428
##
## Relevant coefficients:
## Interaction
##   0.3296193
##
## For standard errors, type coef(summary(x))
```

`parameter_estimates(M1)`

```
##         Estimate    CI95.lo   CI95.hi
## beta  0.9531694 0.8333901 1.0901640
## sigma 0.1034280 0.0000000 0.1176078
```

`Mr`

```
## Stationary Soft core process
## Fitted to point pattern dataset 'Xr'
##
## First order term:  beta = 1.108702
##
## Exponent kappa:  0.2
## Initial approximation to sigma:  0.3079177
## Fitted interaction parameter sigma:  0.303952
##
## Relevant coefficients:
## Interaction
##   0.878415
##
## For standard errors, type coef(summary(x))
```

`parameter_estimates(Mr)`

```
##         Estimate    CI95.lo   CI95.hi
## beta  1.1087018 0.9521537 1.2909889
## sigma 0.3039518 0.2274508 0.3248593
```

Especially the interaction range parameter `sigma` shows a large difference between the point patterns.

Pair correlation functions for all Control subjects in the data

Here the pair correlation functions of all five Control subjects in data show that the analysed pattern exhibits different behavior.

```
dfpcf <- meta |> filter(group=="Controls") |>
  left_join(glands, by="subjectid") |>
  group_by(subjectid) |>
  reframe(pcf(ppp(x=x, y=y, window=owin(c(x0[1], x1[1]), c(y0[1], y1[1]))),
              r=seq(0, 3, length=101)))
dfpcf |> filter(r != 0) |> ggplot() +
  geom_line(aes(x=r, y=trans, group=factor(subjectid), col=factor(subjectid))) +
  theme_minimal() +
  labs(y="pcf") +
  scale_color_discrete(type=grey.colors(2, end=0.7)[c(2,2,2,1,2)]) +
  guides(col="none")
```