Homework 1

Esben Kran

4/6/2022

Statistical signal processing

```
pacman::p_load(tidyverse, ggplot)
```

```
## Installing package into 'C:/Users/esben/Documents/R/win-library/4.0'
## (as 'lib' is unspecified)
```

```
## Warning: package 'ggplot' is not available (for R version 4.0.2)
```

```
## Warning: unable to access index for repository http://www.stats.ox.ac.uk/pub/RWin/bin/wind
ows/contrib/4.0:
##   cannot open URL 'http://www.stats.ox.ac.uk/pub/RWin/bin/windows/contrib/4.0/PACKAGES'
```

```
## Warning in p_install(package, character.only = TRUE, ...):
```

```
## Warning in library(package, lib.loc = lib.loc, character.only = TRUE,
## logical.return = TRUE, : there is no package called 'ggplot'
```

```
## Warning in pacman::p_load(tidyverse, ggplot): Failed to install/load:
## ggplot
```

2.2.1

[NA]

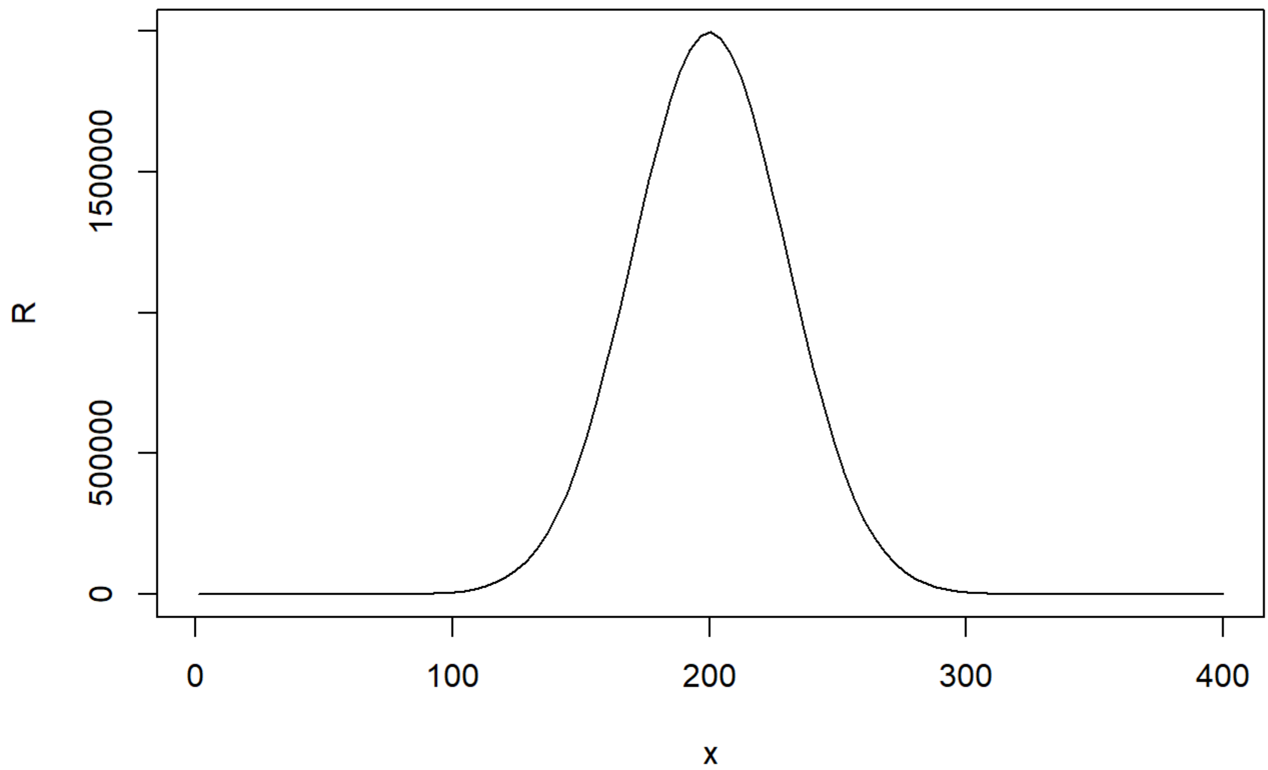
1.1.1

In a radar system an estimator of round trip delay \(\tau_0\) has the PDF \(\hat \tau_0 \sim N(\tau_0,\sigma_{\hat\tau_0}^2)\), where \(\tau_0\) is the true value. If the range is to be estimated, propose an estimator \(R\) and find its PDF. Next determine the standard deviation \(\sigma_{\hat\tau_0}\) so that 99% of the time the range estimate will be within 100 m of the true value. Use \(c=3\cdot 10^8 m/s\) for the speed of electromagnetic propagation.

The estimator for R is \[R~\dfrac{c\cdot N(\tau_0, \sigma^2_{\hat \tau_0})}{2}\]

```
c <- 3 * 10^8
range_lim = 100
sigma = 30
tau = 200

# Defining range estimator:
R <- function(x) (c*dnorm(x, tau, sigma)) / 2
plot(R, xlim=c(1,400))
```




```
# Not sure about determining SD for 99% range estimate \in [-100, 100]
```

1.1.4

It is desired to estimate the value of a DC level A in WGN or \[x[n]=A+w[n]$ where $n=0,1,...,N-1\] Where \(w[n]\) is zero mean and uncorrelated, and each sample has variance \(\sigma^2=1\). Consider the two estimators: \[\hat A = \dfrac{1}{N}\sum^{N-1}_{n=0}{x[n]}\] \[\overline A=\dfrac{1}{N+2}\left(2x[0]+\sum^{N-1}_{n=1}{x[n]+2x[N-1]}\right)\] Which one is better? Does it depend on the value of A?

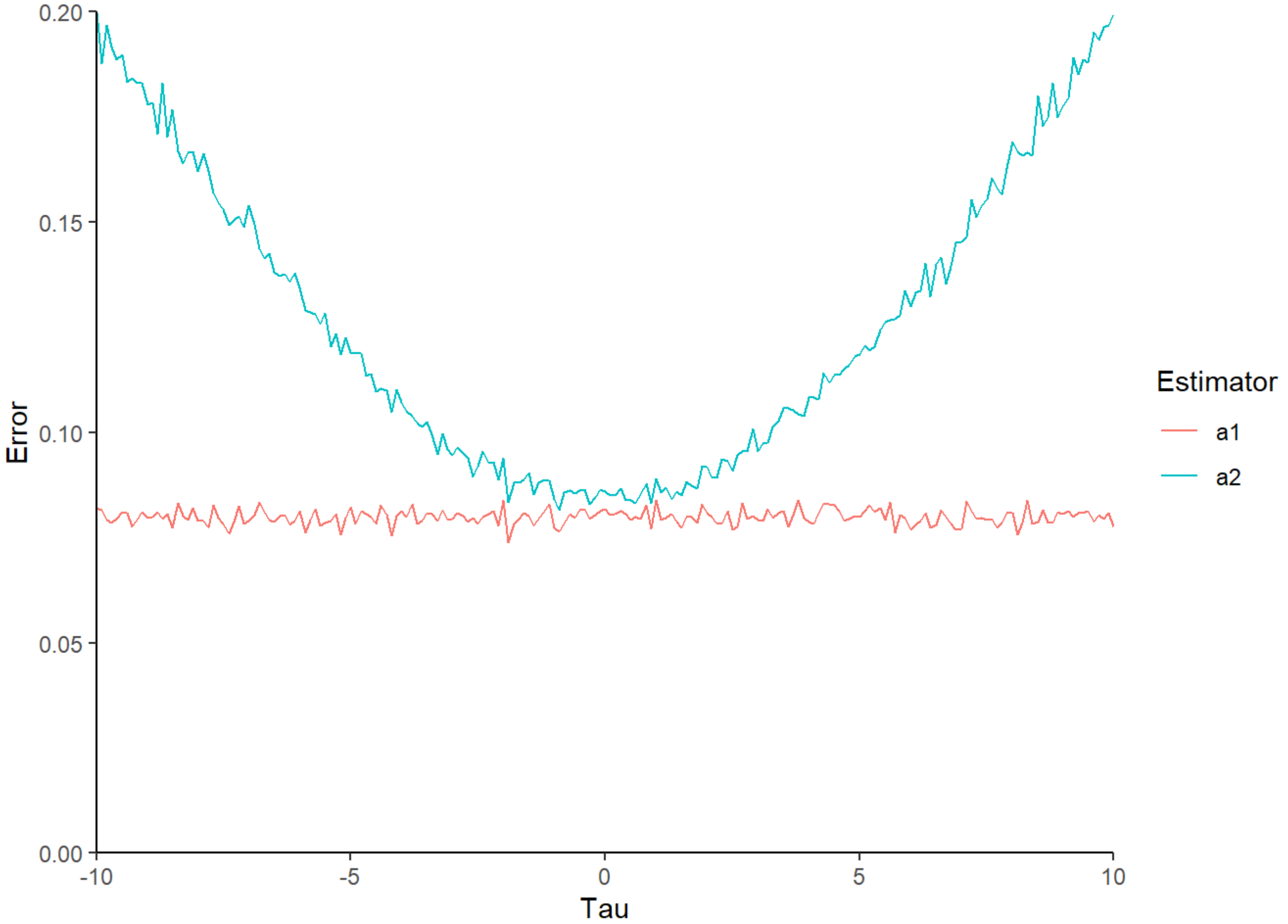
Run M experiments to see difference between the two estimators and analyse error. Based on what we can see here, **the estimator \(A_1\) is always better than estimator \(A_2\)**.

```
test_range = seq(-10, 10, 0.1)
samples = 100
M = 1000

a1 <- function(x) mean(x)
a2 <- function(x) (1/(length(x)+2))*(2*x[1]+sum(x)+2*x[length(x)-1])

data.frame(
    m = lapply(1:M, function(m) rep(m, length(test_range) * samples)) %>% unlist,
    mu = lapply(1:M, function(m) lapply(test_range, function(i) rep(i, samples)) %>% unlist)
 %>% unlist,
    x = lapply(1:M, function(m) lapply(test_range, function(i) rnorm(samples, i, 1)) %>% unli
st) %>% unlist
  ) %>%
  group_by(mu, m) %>%
  summarise(
    a1 = abs(a1(x)-mu),
    a2 = abs(a2(x)-mu)
  ) %>%
  group_by(mu) %>%
  summarise(
    a1 = mean(a1),
    a2 = mean(a2)
  ) %>%
  unique() %>%
  pivot_longer(c(a1, a2)) %>%
  ggplot() +
  aes(mu, value, color = name) +
  geom_line() +
  theme_classic() +
  labs(
    x = "Tau",
    y = "Error",
    color = "Estimator"
  ) +
  coord_cartesian(ylim=c(0, 0.2), expand=F)
```

```
## `summarise()` has grouped output by 'mu', 'm'. You can override using the `.groups` argume
nt.
```



2.1.1

[NA]

1.1.2

An unknown parameter \(O\) influences the outcome of an experiment which is modeled by the random variable x. The PDF of x is \[p(x;O)=\dfrac{1}{\sqrt{2\pi}}exp\left[-\dfrac{1}{2}(x-O)^2\right]\] A series of experiments is performed, and x is found to always be in the interval \([97,103]\). As a result, the investigator concludes that \(O\) must have been 100. Is this assertion correct?

```r
a <- 100
x <- rnorm(1000, 100, 1)
mean_x <- mean(x)
range_x <- max(x) - min(x)

paste(
  "**[1.1.2] Sampling a 1000 times with an A of",
  a,
  "gives a mean of",
  mean_x,
  "with range",
  min(x),
  "to",
  max(x),
  "of",
  range_x,
  "which indicates that the",
  "investigator is correct.**"
)
```

```
## [1] "**[1.1.2] Sampling a 1000 times with an A of 100 gives a mean of 99.9730684183665 wit
h range 96.8803859883147 to 103.187244282681 of 6.30685829436646 which indicates that the inv
estigator is correct.**"
```

6

Suppose \(x\sim N(5,2)\) and \(y\sim 2x+4\). Find \(E(y)\), \(var(y)\), and the PDF \(p_y(y)\).

Expected value (mean(y)):

```r
beta_1 <- function(x)
  dnorm(x, 5, 2)
f <- function(x)
  1/abs(2) * beta_1((x-4)/2)

weighted.mean(seq(-1e2, 1e2, 1e-2), f(seq(-1e2, 1e2, 1e-2)))
```

```
## [1] 14
```

Variance from the sampled distribution:

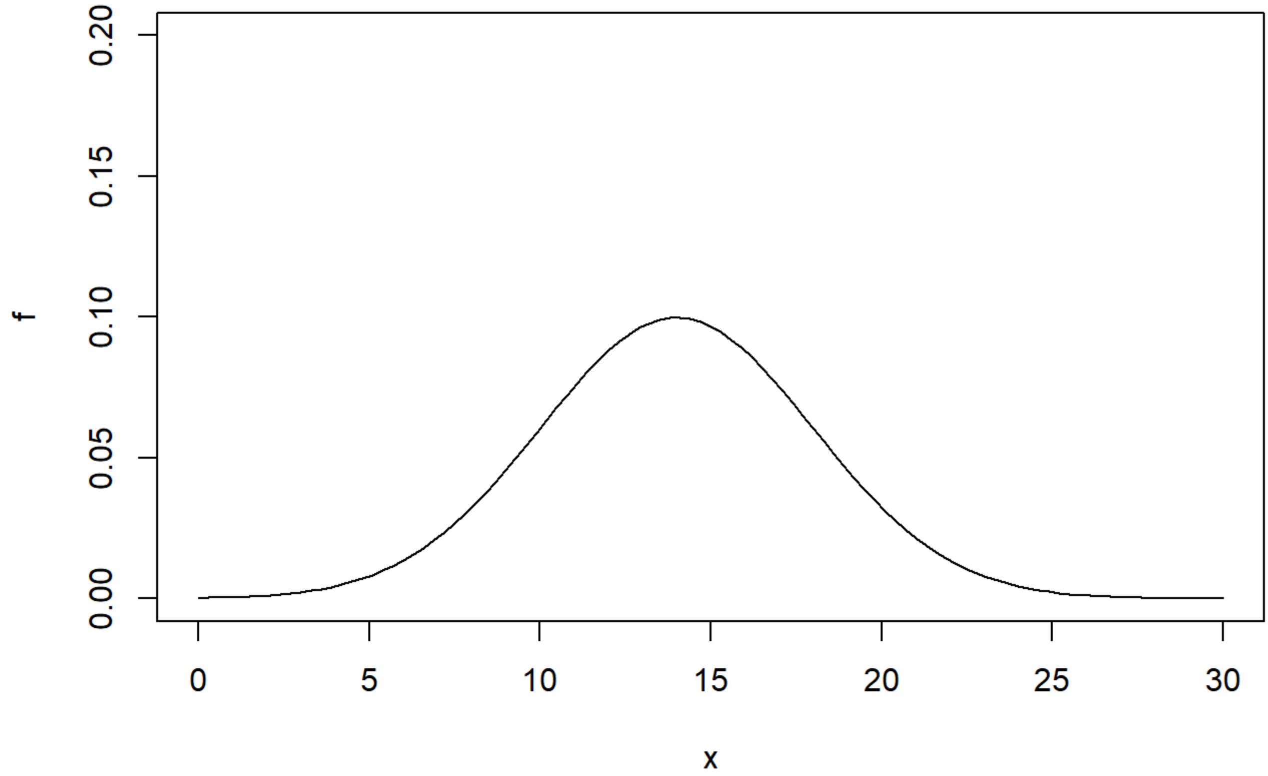
```r
x <- seq(0, 50, 0.1)
px <- f(seq(0, 50, 0.1))
draws <- sample(x, size = 5000, replace = TRUE, prob = px)

var(draws)
```

```
## [1] 15.31712
```

The PDF of y with \(a\) multiplied into \(beta_1\):

```
plot(f, ylim=c(0,0.2), xlim=c(0, 30))
```



7

Suppose that $x\sim N(0,\sigma_x^2)$ and $w\sim N(0,\sigma_w^2)$ and $y=ax+w$. If w and x are independent, what is mean and covariance matrix for the Gaussian vector $z=[x,y]^T$? Hint: Note that $E[z]$ should be a 2D column vector and $var(z)$ should be a 2x2 matrix.

$$z=[x,y]^T$$ $$[E(x), E(\omega)]^T = [0, 0]^T$$ $$var(z)=
\begin{bmatrix}\sigma_x^2&cov(x,y)\\cov(x,y)&var(y)\end{bmatrix}$$