Test different fitting methods

a simple simulation on three different estimates

timing and accuracy

```r
library(sn)
library(ggplot2)
library(dplyr)
library(cowplot)
# mom fit
mom_fit <- function(y){
    n <- length(y)
    # maximum value gamma can take
    max.gamma1 <- 0.5*(4-pi)*(2/(pi-2))^1.5 - (.Machine$double.eps)^(1/4)
    # method of moments strategy
    s <- sd(y)
    gamma1 <- sum((y-mean(y))^3)/(n*s^3)
    if(abs(gamma1) > max.gamma1) gamma1 <- sign(gamma1)*0.9*max.gamma1
    cp1 <- as.numeric(c(mean(y), s, gamma1))
    dp1 <- cp2dp(cp1, family="SN")
    return(dp1)
}

# mle fit
mle_fit <- function(y){
  out <- selm(y~1, method = "MLE")
  return(out@param$dp)
}

# mple fit
mple_fit <- function(y){
  out <- selm(y~1, method = "MPLE")
  return(out@param$dp)
}

# AS fit
as_fit <- function(y){
  # Azzalini and Salehi (2020) strategy
  sn.prelim <- st.prelimFit(y = y, SN=TRUE)
  logL2 <- sn.prelim$logLik
  dp <- sn.prelim$dp
  return(dp)
}
tail_list <- seq(0.9, 0.999, length.out = 500)
no_sam <- round(seq(1e3, 1e6, length.out = 100))
result <- data.frame(no_sam = rep(no_sam, 4),
```

```r
                      acc = 0,
                      timing = 0,
                      type = 0)
for (r in 1:100) {
  set.seed(r)
  data <- rsn(no_sam[r])

  # mom
  start_time <- Sys.time()
  acc <- max(tail_list / psn(quantile(data, tail_list), dp = mom_fit(data)))
  end_time <- Sys.time()
  timing <- as.numeric(end_time - start_time)
  result[r, 2:3] <- c(acc, timing)
  result[r, 4] <- "MoM"

  # MLE
  start_time <- Sys.time()
  acc <- max(tail_list / psn(quantile(data, tail_list), dp = mle_fit(data)))
  end_time <- Sys.time()
  timing <- as.numeric(end_time - start_time)
  result[100 + r, 2:3] <- c(acc, timing)
  result[100 + r, 4] <- "MLE"

  # MPLE
  start_time <- Sys.time()
  acc <- max(tail_list / psn(quantile(data, tail_list), dp = mple_fit(data)))
  end_time <- Sys.time()
  timing <- as.numeric(end_time - start_time)
  result[200 + r, 2:3] <- c(acc, timing)
  result[200 + r, 4] <- "MPLE"

  # AS
  start_time <- Sys.time()
  acc <- max(tail_list / psn(quantile(data, tail_list), dp = as_fit(data)))
  end_time <- Sys.time()
  timing <- as.numeric(end_time - start_time)
  result[300 + r, 2:3] <- c(acc, timing)
  result[300 + r, 4] <- "AS"

}

# plot for accuracy and timing
timing_plt <- result |>
  ggplot(aes_string(x = "no_sam", y = "timing", colour = "type")) +
  geom_line()


acc_plt <- result |>
  ggplot(aes_string(x = "no_sam", y = "acc", colour = "type")) +
  geom_line()

plot_grid(timing_plt,
          acc_plt,
```

```
          ncol = 1,
          align = "v")
```