MSMS 308 : Practical 02

Ananda Biswas

September 11, 2025

Question

Generate random numbers using the Exponential distribution with rate parameter λ. Perform the following tasks:

1. Generate data-sets of sizes 30, 50, and 100 using function in R.
2. For each sample, estimate the rate parameter λ using Maximum Likelihood Estimation.
3. Using the estimated parameter, compute and plot the following functions:
   (a) Probability Density Function $f(t)$
   (b) Survival Function $S(t)$
   (c) Hazard Function $h(t)$
   (d) Cumulative Hazard Function $H(t)$.
4. Compare how the shape and values of these functions change with different sample sizes.

Theory

The probability distribution function of a lifetime T having Exponential Distribution with rate $\lambda > 0$ is

$$f(t) = \lambda e^{-\lambda t} \cdot I_{(0,\infty)}(t).$$

The Survival Function is $S(t) = e^{-\lambda t}\ \forall t > 0$.

The Hazard Function is $h(t) = \lambda\ \forall t > 0$.

The Cumulative Hazard Function is $H(t) = \lambda t\ \forall t > 0$.

By the method of Maximum Likelihood Estimation, $\hat{\lambda} = \dfrac{1}{\overline{T}}$.

R Program

```
f_t <- function(t, lambda) lambda * exp(-lambda * t)
S_t <- function(t, lambda) exp(-lambda * t)
h_t <- function(t, lambda) lambda
H_t <- function(t, lambda) lambda * t
```

⚀ $n = 30$.

```r
n <- 30
```

```r
s <- rexp(n, rate = 0.5)
```

```r
lambda_hat <- 1 / mean(s); lambda_hat

## [1] 0.4611669
```

```r
df1 <- data.frame(t = s,
                  f_t_hat = f_t(s, lambda_hat),
                  S_t_hat = S_t(s, lambda_hat),
                  h_t_hat = h_t(s, lambda_hat),
                  H_t_hat = H_t(s, lambda_hat))
```

⚁ $n = 50$.

```r
n <- 50
```

```r
s <- rexp(n, rate = 0.5)
```

```r
lambda_hat <- 1 / mean(s); lambda_hat

## [1] 0.4625597
```

```r
df2 <- data.frame(t = s,
                  f_t_hat = f_t(s, lambda_hat),
                  S_t_hat = S_t(s, lambda_hat),
                  h_t_hat = h_t(s, lambda_hat),
                  H_t_hat = H_t(s, lambda_hat))
```

⚂ $n = 100$.

```r
n <- 100
```

```r
s <- rexp(n, rate = 0.5)
```

```r
lambda_hat <- 1 / mean(s); lambda_hat

## [1] 0.5050291
```

```r
df3 <- data.frame(t = s,
                  f_t_hat = f_t(s, lambda_hat),
                  S_t_hat = S_t(s, lambda_hat),
                  h_t_hat = h_t(s, lambda_hat),
                  H_t_hat = H_t(s, lambda_hat))
```

➜ Plots

⊡ $n = 30$.

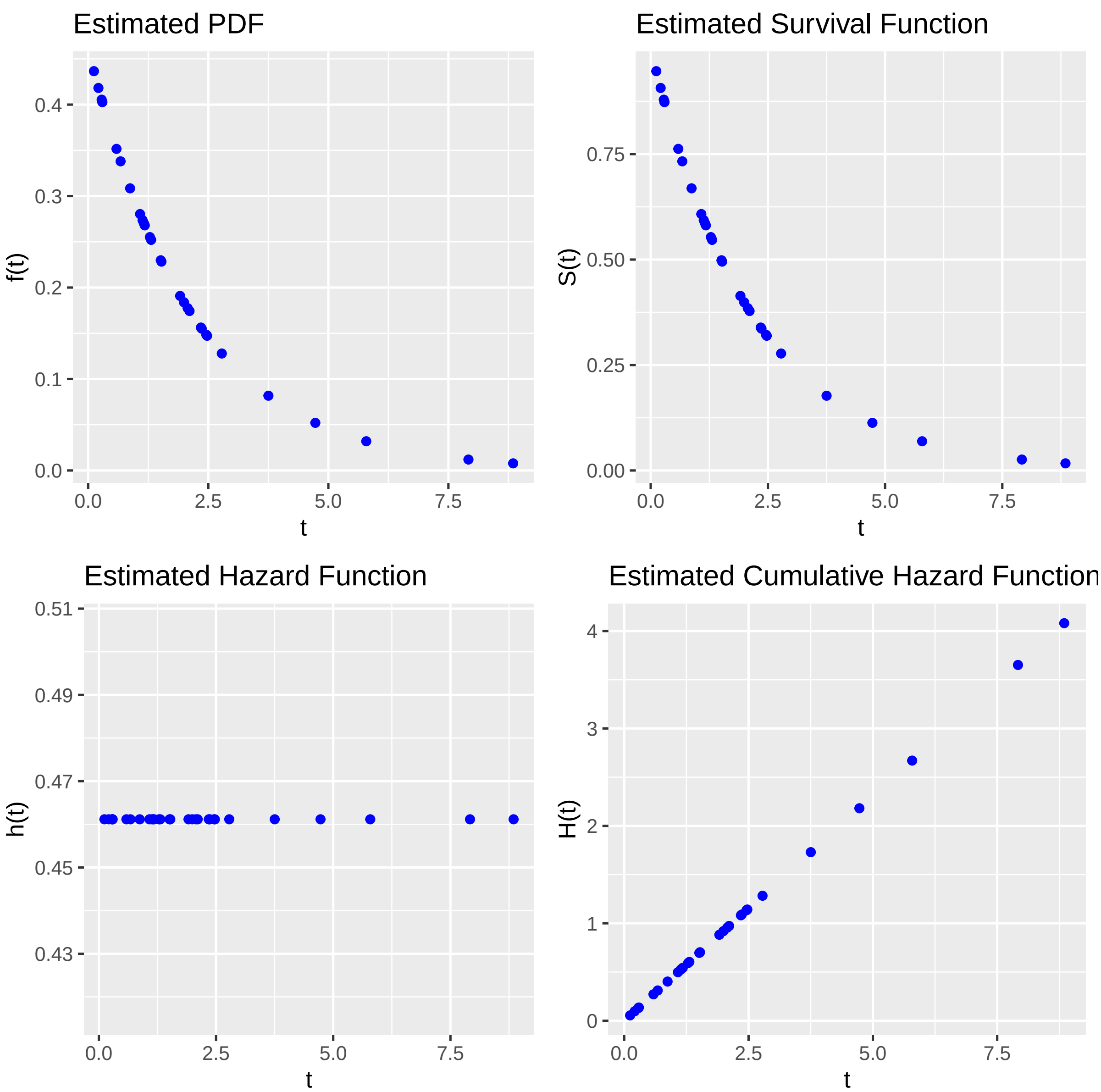
```
p1 <- df1 %>%
        ggplot(aes(x = t, y = f_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "f(t)", title = "Estimated PDF")

p2 <- df1 %>%
        ggplot(aes(x = t, y = S_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "S(t)", title = "Estimated Survival Function")

p3 <- df1 %>%
        ggplot(aes(x = t, y = h_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "h(t)", title = "Estimated Hazard Function")

p4 <- df1 %>%
        ggplot(aes(x = t, y = H_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "H(t)", title = "Estimated Cumulative Hazard Function")
```

```
grid.arrange(p1, p2, p3, p4, ncol = 2)
```



- $n = 50$.

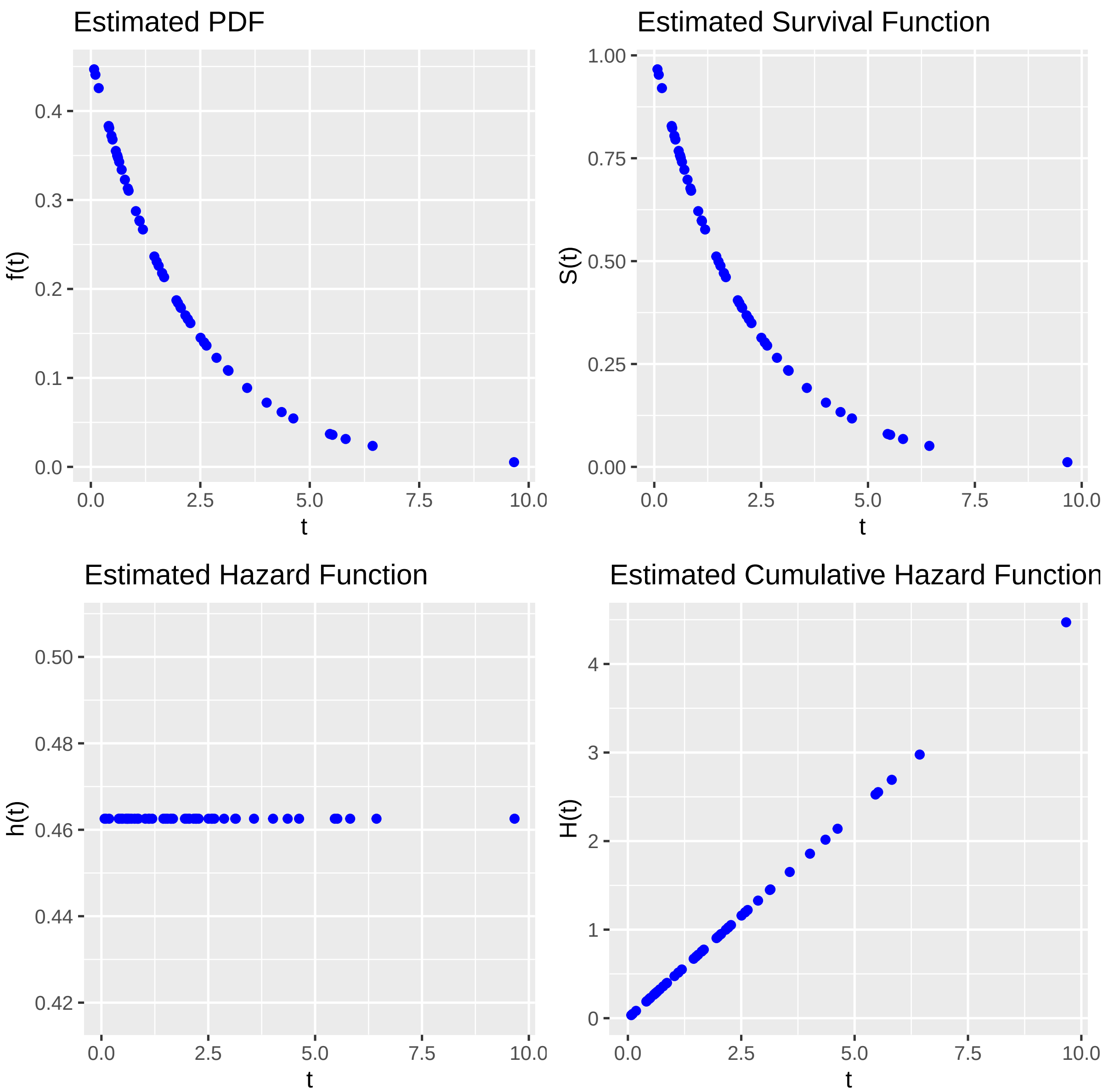
```
p1 <- df2 %>%
        ggplot(aes(x = t, y = f_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "f(t)", title = "Estimated PDF")

p2 <- df2 %>%
        ggplot(aes(x = t, y = S_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "S(t)", title = "Estimated Survival Function")

p3 <- df2 %>%
        ggplot(aes(x = t, y = h_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "h(t)", title = "Estimated Hazard Function")

p4 <- df2 %>%
        ggplot(aes(x = t, y = H_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "H(t)", title = "Estimated Cumulative Hazard Function")
```

```
grid.arrange(p1, p2, p3, p4, ncol = 2)
```



⚁ $n = 100$.

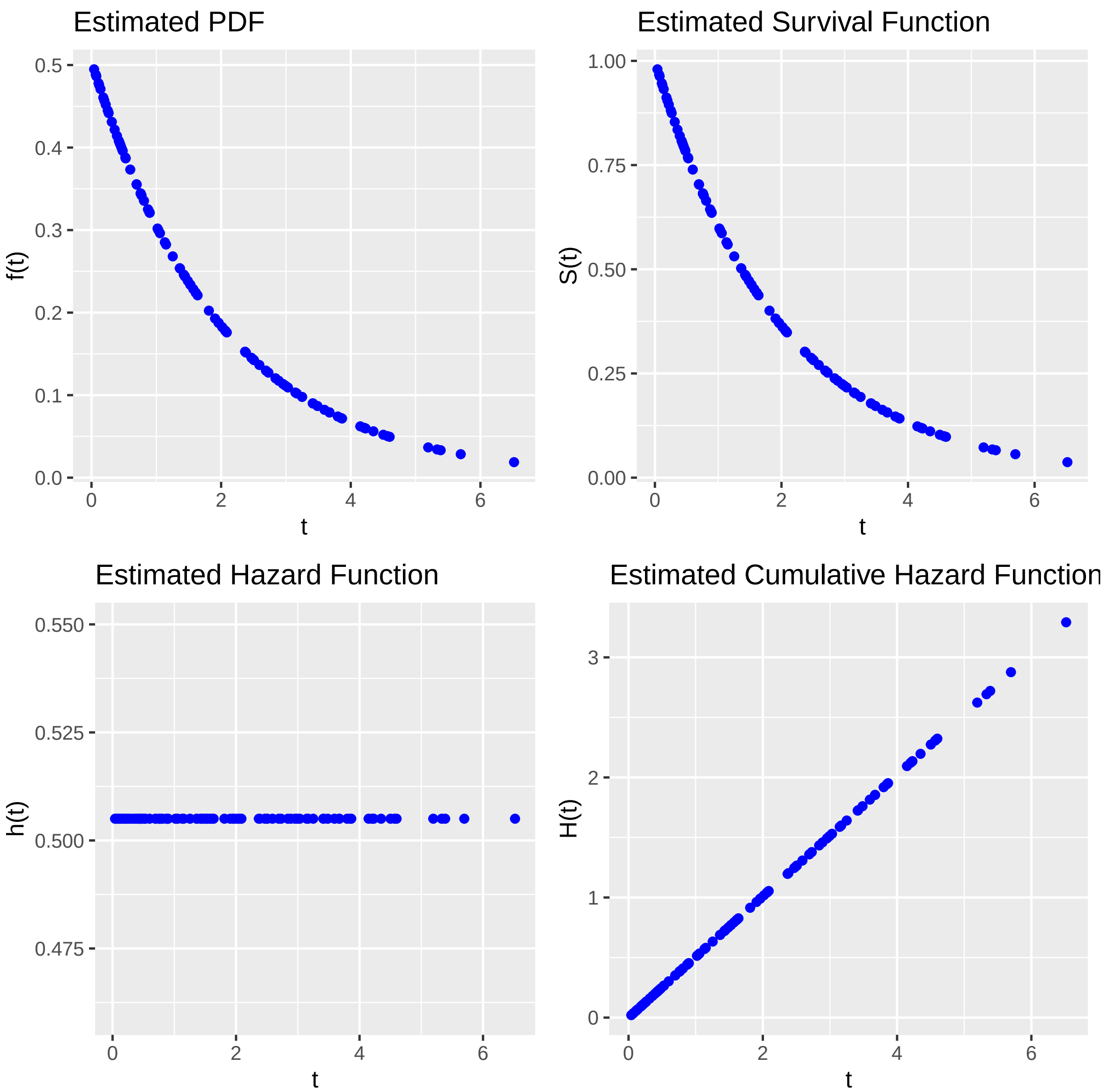
```
p1 <- df3 %>%
        ggplot(aes(x = t, y = f_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "f(t)", title = "Estimated PDF")

p2 <- df3 %>%
        ggplot(aes(x = t, y = S_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "S(t)", title = "Estimated Survival Function")

p3 <- df3 %>%
        ggplot(aes(x = t, y = h_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "h(t)", title = "Estimated Hazard Function")

p4 <- df3 %>%
        ggplot(aes(x = t, y = H_t_hat)) + geom_point(col = "blue") +
        labs(x = "t", y = "H(t)", title = "Estimated Cumulative Hazard Function")
```

```
grid.arrange(p1, p2, p3, p4, ncol = 2)
```



➔ Conclusion

As the sample size increases, the ML estimate of λ becomes less noisy. Consequently the estimated PDF, Survival Function, Hazard Function and Cumulative Hazard match with true shapes of these functions more closely.