Statistical Inference: The Study of the Exponential Distribution, A Simulation Exercise

Alexander Tuzhikov

September 14, 2015

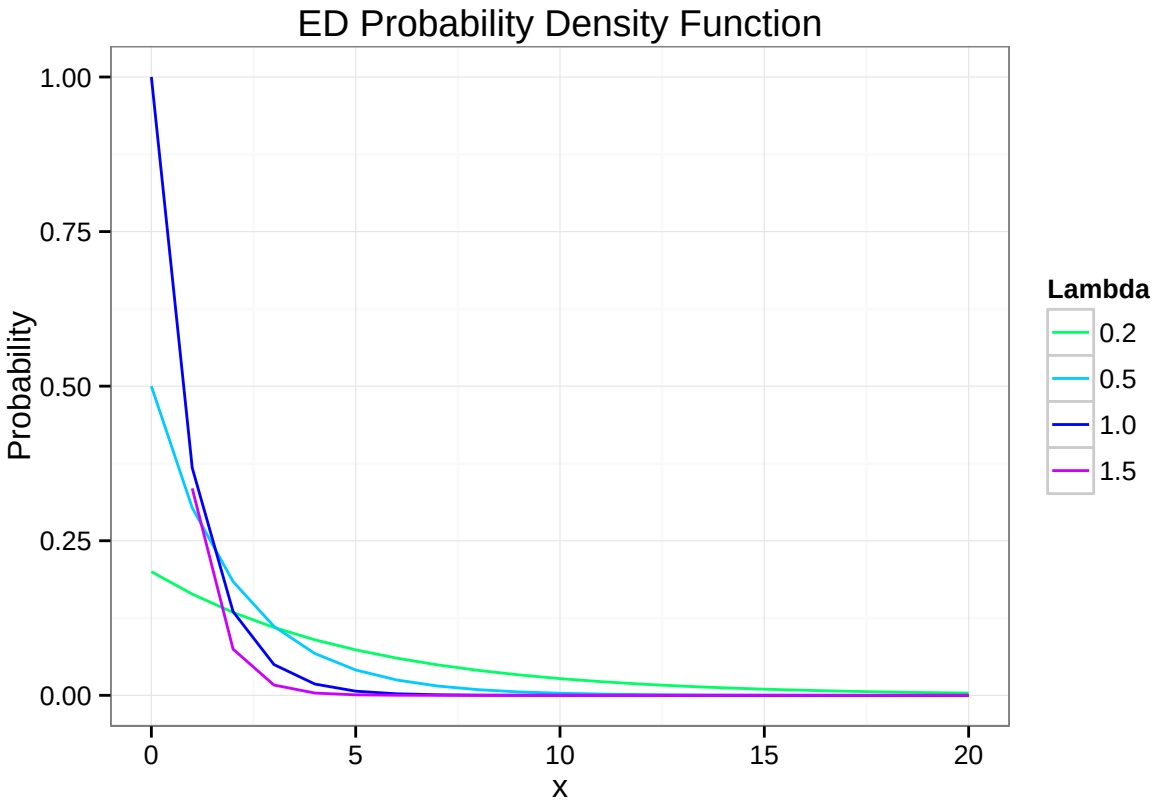
Contents

1 Overview: Exponential Distribution

In accordance with Wikipedia, **e**xponential **d**istribution (ED) *is the probability distribution that describes the time between events in a Poisson process, i.e. a process in which events occur continuously and independently at a constant average rate.* Both *mean* and *standard deviation* of the ED is *1/lambda*. As suggested in the study objective, here we will use lambda = 0.2. However, for the purpose of introduction, let's reconstruct the wiki plots of the ED with different lambda:

```
lambdas<- c(0.2, 0.5, 1, 1.5)#the given in the task + those from wikipedia
n<- 40 #given by ".. you will investigate the distribution of
       #averages of 40 exponential(0.2)s" in the task
sampling.count<- 1000  #given by ".. you will need to do a
       #thousand or so simulated averages of 40 exponentials" in the task
#prepare a data.frame for the plot, melt by x, plot as line
ed.plot.df<- as.data.frame(cbind(
        x=0:40,
        la.0.2=dexp(x=0:40, lambdas[1]),
        la.0.5=dexp(x=0:40, lambdas[2]),
        la.1=dexp(x=0:40, lambdas[3]),
        la.1.5=dexp(x=0:40, lambdas[4])
)) %>%
        melt(id.vars="x") %>%
        ggplot(data=., mapping=aes(x=x, group=variable, y=value, color=variable)) +
        geom_line() + theme_bw() + xlim(0,20) + ylim(0,1) +
        labs(title="ED Probability Density Function") + ylab("Probability") +
        scale_color_manual(values=rainbow(4, start = 0.4, end = 0.8), labels=c("0.2","0.5", "1.0", "1.5
plot(ed.plot.df)
```

```
## Warning: Removed 81 rows containing missing values (geom_path).
```



2 Simulations

3 Sample Mean versus Theoretical Mean

4 Sample Variance versus Theoretical Variance

5 Distribution