Statistical Inference Final Project (Part 1)

MP

Sunday, September 27, 2015

Introduction

The basic concept behind Central Limit Theorem(CLT) is that as increase occurs in number of iid samples mean values of samples converges to standard normal state so the mean and standard deviation characteristic converges to population mean and standard deviation. in this report theoretical concept of Exponenntial distribution compared with results of Central Limit Theorem.

In probability theory and statistics[1], the exponential distribution (a.k.a. negative exponential distribution) 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. It is a particular case of gamma distribution. based on theory the expected value is $E[x] = 1/\lambda$ and standard deviation of exponentioal distribution is $\sigma[x] = 1/\lambda$.

[1] http://en.wikipedia.org/wiki/Exponential_distribution

```
set.seed(3)
lambda <- 0.2
sample <- 40
for (n in c(sample)) {
    temp <- matrix(rexp(n * 1000, rate = lambda),
        ncol = n)
    temp <- apply(temp, 1, mean)
    hist(temp, breaks = 50, prob = TRUE, main = "sample average with lambda=0.2",
        xlab = "", col = "coral3")
    dty <- density(temp)
    lines(dty, col = "red", lty = 1)
    xfit <- seq(range(temp)[1], range(temp)[2],
        length = length(dty$x))
    yfit <- dnorm(xfit, mean = 1/lambda, sd = (1/lambda/sqrt(sample)))
    lines(xfit, yfit, pch = 22, col = "blue",
        lty = 2)
    abline(v = mean(temp), col = "green", lwd = 3,
        lty = 3)
    abline(v = 1/lambda, col = "brown", lwd = 3,
        lty = 4)
    legend("topright", c("simulation", "theoretical",
        "simulation average line", ... = "theoretical average line"),
        lty = c(1, 2, 3, 4), col = c("red", "blue",
```

```
            "green", "brown"), title.adj = 10)
}
```

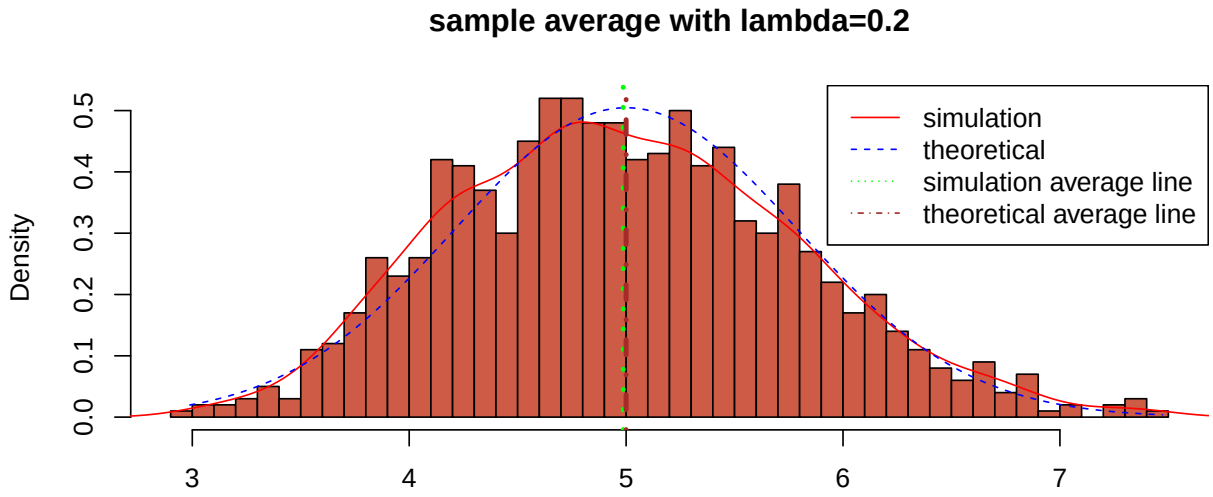

Figure 1: Comparison between Simulation and theoretical of Exponential distribution

In 1000 times of simulation the constructed sample average curve is with mean of about 5.011 that in theory should be 5 and standard deviation of about 0.780 that in theory should be about 0.790.

For evaluating the variation of sample normal distribution the theoretical normal values Mean squered error in curve structure have been used. here dty$y is the output of simulation result and yfit is the theoretical output Value.

```
sum(((dty$y - yfit)^2)/length(yfit))
```

```
## [1] 0.002177917
```

$$MSE = \frac{\sum_{n=0}^{\infty} (output - target)^2}{N}$$

Figure 2: MSE equation