Math3810 - Probability
Section 001 - Fall 2025
Notes: More Expectation Through Conditioning

University of Colorado Denver / College of Liberal Arts and Sciences

Department of Mathematics - Dr. Robert Rostermundt

The Problem:

A professor has accidentally scheduled two different appointments at the same time. One student show up on time and the other student shows up 5 minutes late. Assume that both appointment times are independent and each is modeled by an exponential distribution with average time 30 minutes. If T is the elapsed time from the first student arrival until the end of the second appointment, find the expectation $E[T]$.

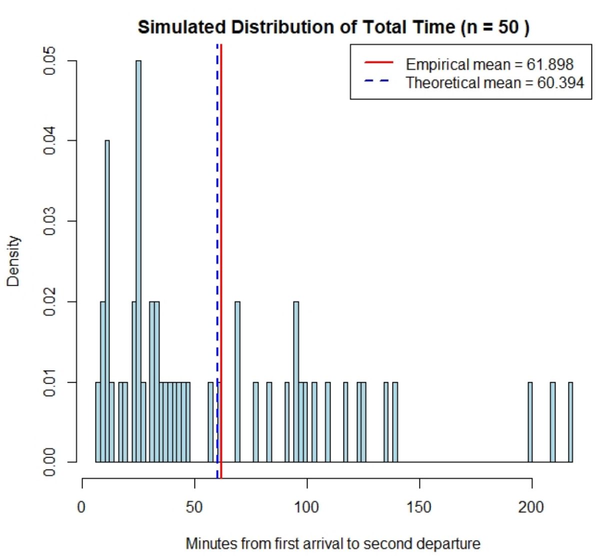
Let X_1 be the time of the first appointment and X_2 be the time of the second appointment. Then we have $X_1 \sim exp\big(\lambda = 1/30\big)$ and $X_2 \sim exp\big(\lambda = 1/30\big)$ and $X_1 \perp X_2$. We want to find

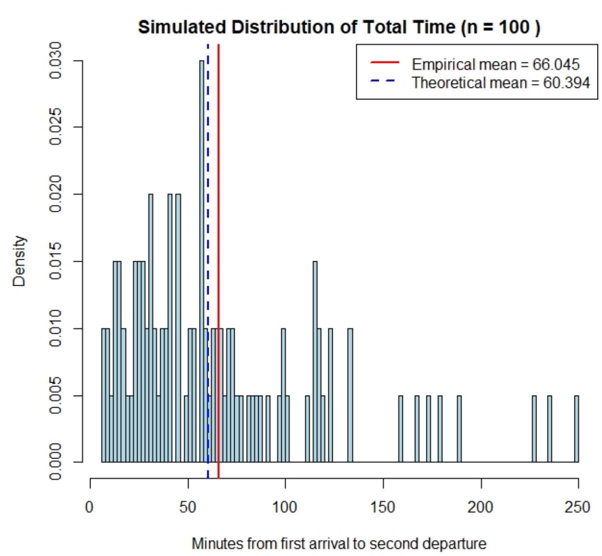
$$E[T] = \int_{-\infty}^{\infty} t \cdot f_T(t)\, dt.$$

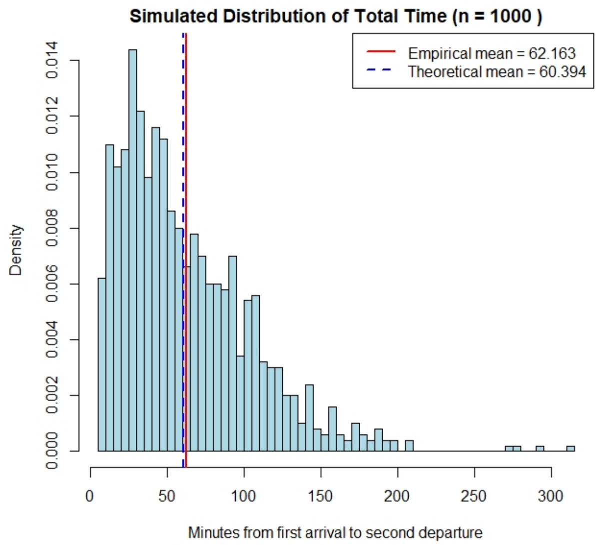
However, at this point we do not know the distribution of T and do not know it's density function f_T. Before we make any theoretical calculations lets simulate the process and see if we can gain any insight.

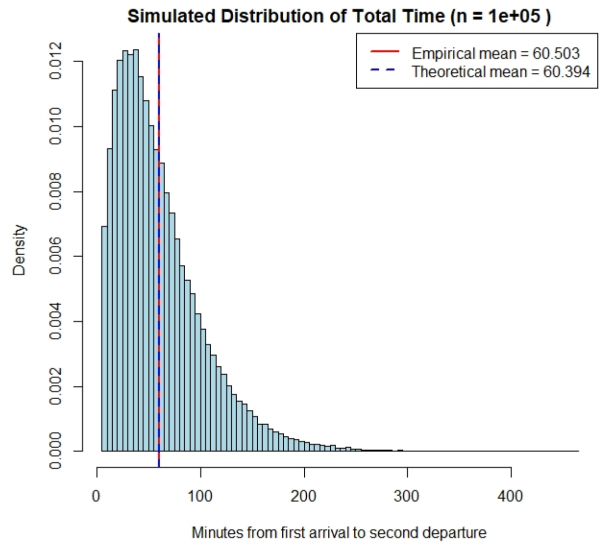
Simulations:

I have simulated this process in R. Below are graphics from four different simulations with number trials $n = 50, 100, 1000, 100000$. The R code is provided the end of the document.



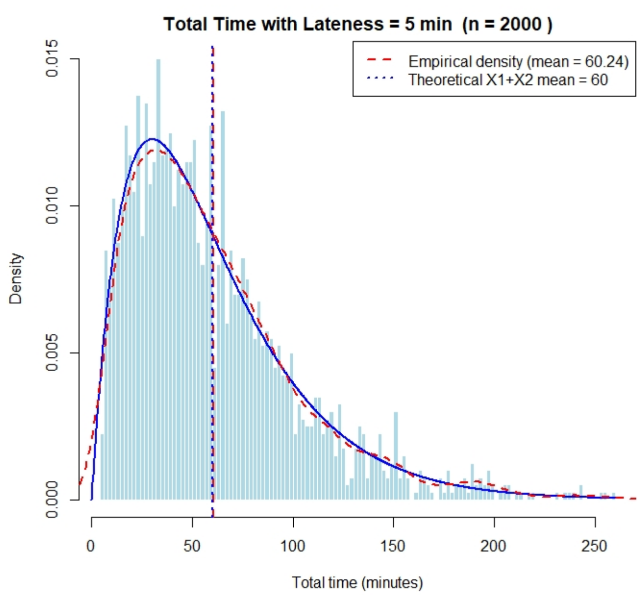


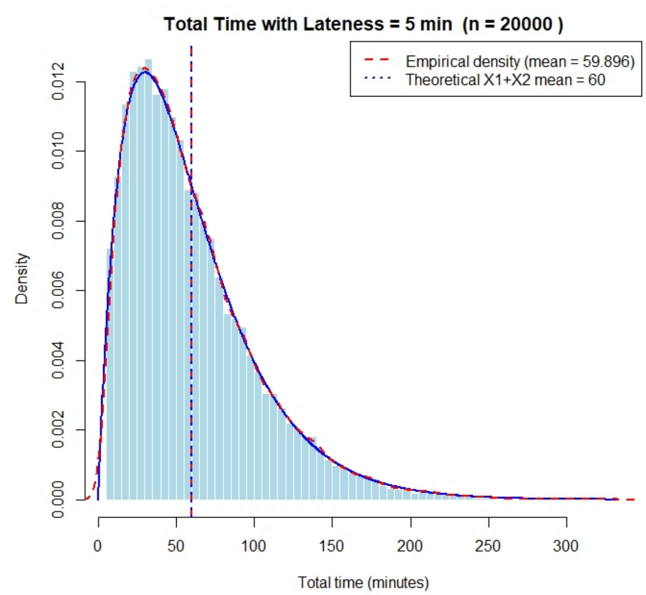




Each simulation is labeled with the true theoretical mean, but without this theoretical information these simulations might give us a sense for the desired expected value $E[T]$.

Next are a couple of comparisons of our simulations of the density function f_T with the density function of the sum $X_1 + X_2$. Here we can see how the late arrival is affecting the expected elapsed time.





Theoretical Tools:

We will compute this expectation in two ways by conditioning on the following events:

> $A := \{0 < X_1 < 5\}$ which is the event that the first appointment lasts less than 5 minutes.
>
> $B := \{5 \le X_1\}$ which is the event that the first appointment lasts at least 5 minutes.

Then since A and B partition the sample space we can compute the expected value as

$$E[T] = \mathbb{P}(A) \cdot E\big[\,T \,|\, A\,\big] + \mathbb{P}(B) \cdot E\big[\,T \,|\, B\,\big].$$

Calculations:

We first make the easy calculations

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(0 < X_1 \le 5) = F_{X_1}(5) = 1 - e^{-5/30} \\ \mathbb{P}(B) &= 1 - \mathbb{P}(A) = e^{-5/30} \end{aligned}$$

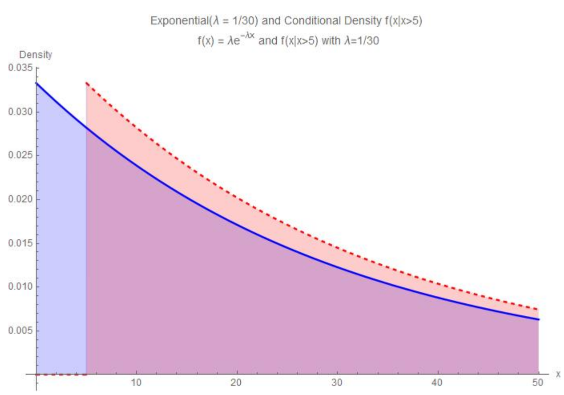
We need to find $E\big[\,T \,|\, A\,\big]$ and $E\big[\,T \,|\, B\,\big]$. So

$$E\big[\,T \,|\, A\,\big] = E\big[\,5 + X_2 \,\big|\, 0 < X_1 < 5\,\big] = 5 + E\big[\,X_2 \,\big|\, 0 < X_1 < 5\,\big] = 5 + E[X_2] = 35.$$

Similarly,

$$E\big[\,T \,|\, B\,\big] = E\big[\,X_1 + X_2 \,\big|\, 5 \le X_1\,\big] = E\big[\,X_1 \,\big|\, 5 \ge X_1\,\big] + E\big[\,X_2 \,\big|\, 5 \ge X_1\,\big] = 5 + E[X_1] + E[X_2] = 65.$$

To see this, recall that the exponential random variable X_1 is memoryless. The following plot should help to see that the conditional distribution $X_1 \,|\, X_1 > 5$ is identical to the distribution of X_1 but with all mass shifted to the right 5 units.



Putting the pieces together we have

$$E[T] = \underbrace{\mathbb{P}(A)}_{1 - e^{-5/30}} \cdot \underbrace{E\big[\,T \,|\, A\,\big]}_{35} + \underbrace{\mathbb{P}(B)}_{e^{-5/30}} \cdot \underbrace{E\big[\,W \,|\, B\,\big]}_{65}$$

So the expected waiting time is 60.394 minutes.

The R Code: Here is the R code used for the above simulations.

```
# Simulated density and expectation of total elapsed time T

set.seed(123)
# parameters
n <- 100000           # number of simulated trials
lambda <- 1/30        # Exp rate, mean = 30 minutes



# simulate appointment durations
x1 <- rexp(n, rate = lambda)
x2 <- rexp(n, rate = lambda)

# second student arrives 5 minutes late
lateness <- 5

# compute relevant times for each trial
start1 <- rep(0, n)
end1   <- x1
arrive2 <- rep(lateness, n)
start2  <- pmax(end1, arrive2)
end2    <- start2 + x2
T <- end2   # total time from first arrival to second departure

# empirical mean
mean_T <- mean(T)

# print summary
cat("Estimated E[T] =", mean_T, "minutes\n")

# per-trial comparison (first 10)
results <- data.frame(
  trial = 1:10,
  X1 = round(x1[1:10], 2),
  X2 = round(x2[1:10], 2),
  end1 = round(end1[1:10], 2),
  start2 = round(start2[1:10], 2),
  end2 = round(end2[1:10], 2),
  Total = round(T[1:10], 2)
)
print(results)



# summary statistics
cat("\nMean(X1) =", mean(x1), "  Mean(X2) =", mean(x2), "\n")
cat("Mean(max(X1,5)) =", mean(pmax(x1,5)), "\n")
cat("Mean(Total)      =", mean_T, "\n")

# plot with n and empirical mean in legend
hist(T, breaks = 100, freq = FALSE,  col = "lightblue",
```

```r
     main = paste("Simulated Distribution of Total Time (n =", n, ")"),
     xlab = "Minutes from first arrival to second departure")

abline(v = mean_T, col = "red", lwd = 2)
abline(v = 60.394, col = "blue", lwd = 2, lty = 2)

legend("topright",
       legend = c(
         paste0("Empirical mean = ", round(mean_T, 3)),
         "Theoretical mean = 60.394"
       ),
       col = c("red", "blue"),
       lwd = 2,
       lty = c(1, 2))

########################################################################
########################################################################
# Same simulation except comparison to Density of X1+X2 without Late Arrival


set.seed(123)

# parameters
n <- 200000
lambda <- 1/30      # Exp(mean=30)
lateness <- 5       # second student 5 minutes late

# simulate appointment durations
x1 <- rexp(n, rate = lambda)
x2 <- rexp(n, rate = lambda)

# total time from first arrival (0) to second departure
T <- pmax(x1, lateness) + x2

# compute means
emp_mean <- mean(T)
theory_mean <- 2 / lambda    # mean of X1 + X2

# histogram of simulated totals
hist(T, breaks = 100, freq = FALSE, col = "lightblue",
     main = paste("Total Time with Lateness =", lateness, "min  (n =", n, ")"),
     xlab = "Total time (minutes)", border = "white")

# overlay theoretical Gamma(2, rate=lambda) density
tseq <- seq(0, max(T), length.out = 1000)
lines(tseq, dgamma(tseq, shape = 2, rate = lambda), col = "blue", lwd = 2)

# overlay empirical density curve
lines(density(T), col = "red", lwd = 2, lty = 2)

# add vertical lines for means
abline(v = emp_mean, col = "red", lwd = 2, lty = 2)
abline(v = theory_mean, col = "blue", lwd = 2, lty = 3)
```

```
# legend
legend("topright",
       legend = c(
         paste0("Empirical density (mean = ", round(emp_mean, 3), ")"),
         paste0("Theoretical X1+X2 mean = ", theory_mean)
       ),
       col = c("red", "blue"),
       lwd = 2,
       lty = c(2, 3),
       bg = "white")
```

Please let me know if you have any questions, comments, or corrections!