Sistem Antrian pada Kasir Kantin BKL dengan Model M/G/1

Import Data

Mengimport data mentah yang diambil dari lapangan untuk dianalisis

```
{r}
data <- read_csv("college/SZN 7/PemStok/Tubes/Dataset_BKL.csv")
View(data)
```

```
Rows: 81 Columns: 4
── Column specification ──────────────────────────────────────────────────────
Delimiter: ","
dbl  (1): id
time (3): wk, wmd, wsd

ℹ Use `spec()` to retrieve the full column specification for this data.
ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Preprocessing

menghapus jika terdapat baris berisi nilai kosong, data asli dengan ukuran 81x4 setelah pengahpusan tetap 81x4 berarti tidak ada missing value

```
{r}
data <- na.omit(data)
data
```

A tibble: 81 x 4

id <dbl>	wk <S3: hms>	wmd <S3: hms>	wsd <S3: hms>
1	43215 secs	43215 secs	43271 secs
2	43268 secs	43272 secs	43310 secs
3	43306 secs	43312 secs	43355 secs
4	43403 secs	43403 secs	43432 secs
5	43557 secs	43558 secs	43609 secs
6	43600 secs	43611 secs	43648 secs
7	43655 secs	43655 secs	43670 secs
8	43668 secs	43672 secs	43697 secs
9	43670 secs	43698 secs	43712 secs
10	43697 secs	43712 secs	43724 secs

1-10 of 81 rows   Previous 1 2 3 4 5 6 ... 9 Next

memastikan format waktu yang dapat dibaca oleh R

```
{r}
data$wk <- as.POSIXct(data$wk, format = "%H:%M:%S")
data$wmd     <- as.POSIXct(data$wmd, format = "%H:%M:%S")
data$wsd   <- as.POSIXct(data$wsd, format = "%H:%M:%S")
```

Karena data sudah cukup bersih dari CSV-nya tidak dilakukan preprocessing berlebihan

Hitung interarrival & service time

Proses ini untuk menghitung dua komponen penting dalam analisis sistem antrian, yaitu **service time** dan **interarrival time**

```
{r}
# service time (detik)
data$service_time <- as.numeric(difftime(data$wsd, data$wmd, units="secs"))

# interarrival time (detik)
data <- data[order(data$wk), ]
data$interarrival <- c(NA, diff(data$wk))
data$interarrival <- as.numeric(data$interarrival, units="secs")
```

```
{r}
summary(data$service_time)
summary(data$interarrival)
```

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 12.0    26.0    31.0    31.4    36.0    59.0
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 2.00   14.00   29.50   44.59   62.00  229.00       1
```

Dari hasil *summary* pada variabel **service time**, diperoleh nilai rata rata sebesar **31.4 detik**. Nilai ini menggambarkan waktu pelayanan rata rata yang dibutuhkan petugas untuk melayani satu pelanggan. Sementara itu, dari hasil *summary* pada variabel **interarrival time**, diperoleh nilai rata rata sebesar **44.59 detik**. Nilai ini menunjukkan rata rata waktu antar kedatangan pelanggan. Ditemukan satu nilai **NA** pada interarrival time, yang merupakan kondisi normal karena pelanggan pertama tidak memiliki pelanggan sebelumnya untuk dihitung selisih kedatangannya.

Estimasi parameter

```
{r}
# Hitung mean dan variansi service time
mean_service <- mean(data$service_time)
var_service  <- var(data$service_time)
# Hitung mean interarrival
mean_inter <- mean(data$interarrival, na.rm = TRUE)
# Parameter M/G/1
lambda <- 1 / mean_inter          # laju kedatangan
ES     <- mean_service            # rata rata service time
VarS   <- var_service             # variansi service time
rho <- lambda * ES                # utilisasi
lambda; ES; VarS; rho
```

```
[1] 0.02242781
[1] 31.39506
[1] 84.84198
[1] 0.7041225
```

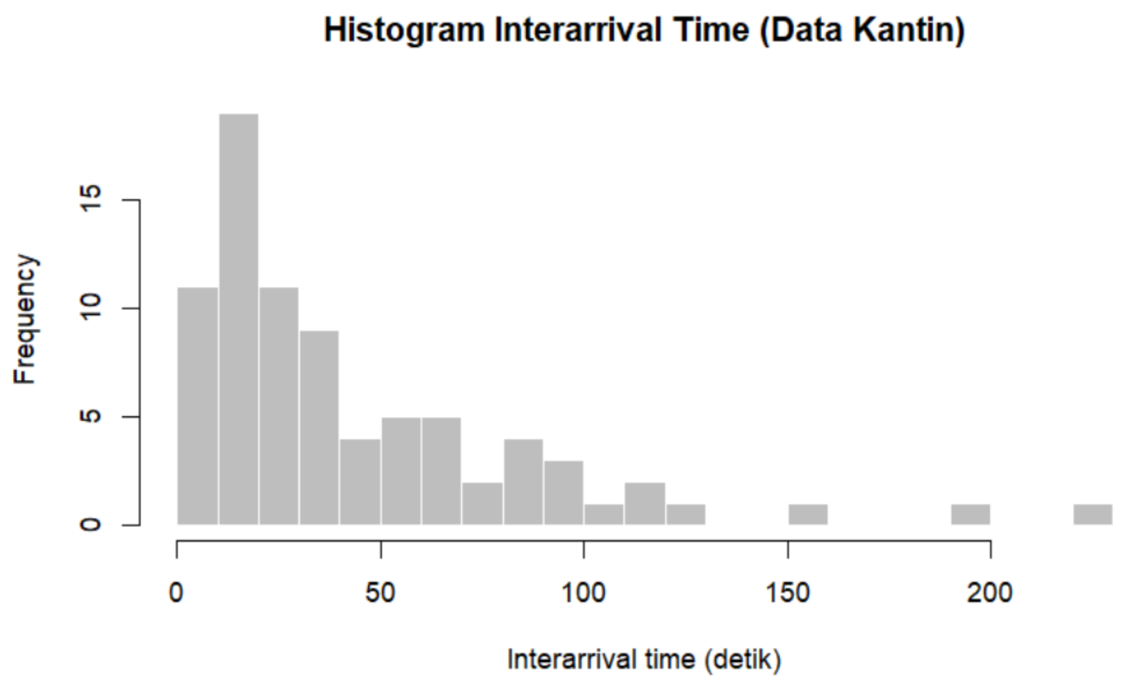
Berdasarkan hasil perhitungan, waktu antar kedatangan rata rata pada data adalah sekitar **44.59 detik**, sehingga laju kedatangan pelanggan menjadi **λ = 0.02243 kedatangan per detik**. Sementara itu, waktu pelayanan memiliki rata rata sebesar **31.40 detik** dengan variansi **84.84**. Dengan parameter tersebut, utilisasi sistem dihitung sebagai **ρ = λ E[S] ≈ 0.704**, yang berarti petugas aktif melayani pelanggan sekitar 70 persen dari waktu operasionalnya. Nilai ini masih berada di bawah batas kestabilan sistem ($\rho < 1$), sehingga proses antrian tetap berada pada kondisi stabil dan dapat digunakan untuk perhitungan performa teori model M/G/1 pada tahap selanjutnya.

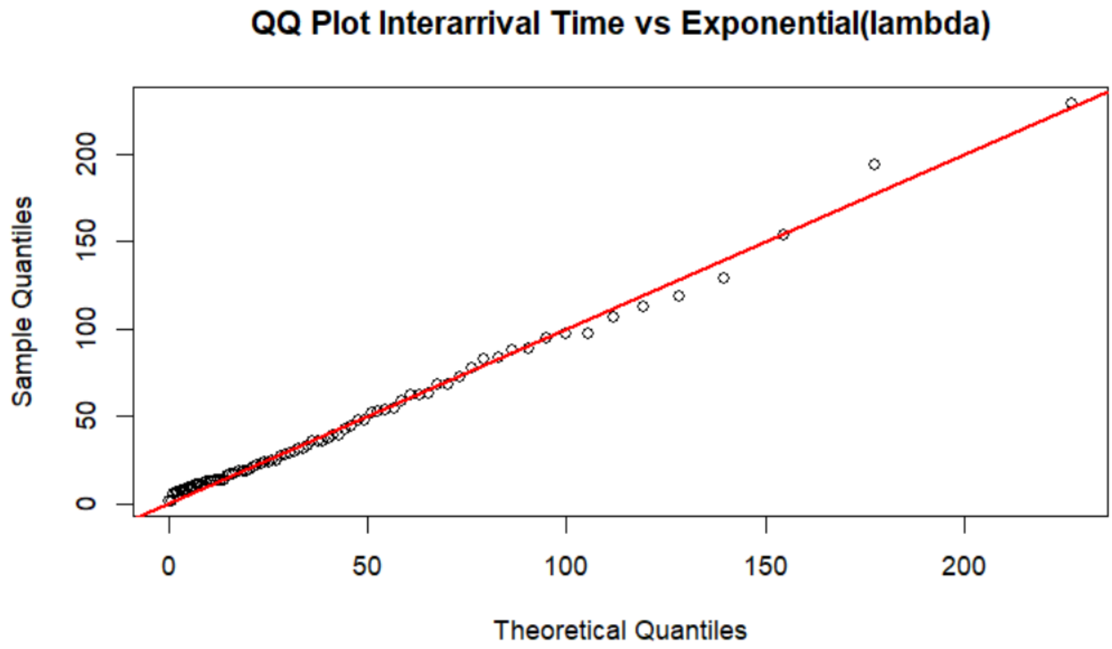
Validasi asumsi distribusi

untuk memvalidasi model apa yang tepat digunakan kita cek terlebih dahulu distribusinya

>Validasi 1: interarrival time

```
{r}
# Histogram interarrival
hist(
  data$interarrival,
  breaks = 30,
  main = "Histogram Interarrival Time (Data Kantin)",
  xlab = "Interarrival time (detik)",
  col = "gray",
  border = "white"
)
# QQ plot interarrival vs eksponensial(lambda)
qqplot(
  qexp(ppoints(length(data$interarrival)), rate = lambda),
  data$interarrival,
  main = "QQ Plot Interarrival Time vs Exponential(lambda)",
  xlab = "Theoretical Quantiles",
  ylab = "Sample Quantiles"
)
abline(0, 1, col = "red", lwd = 2)
```

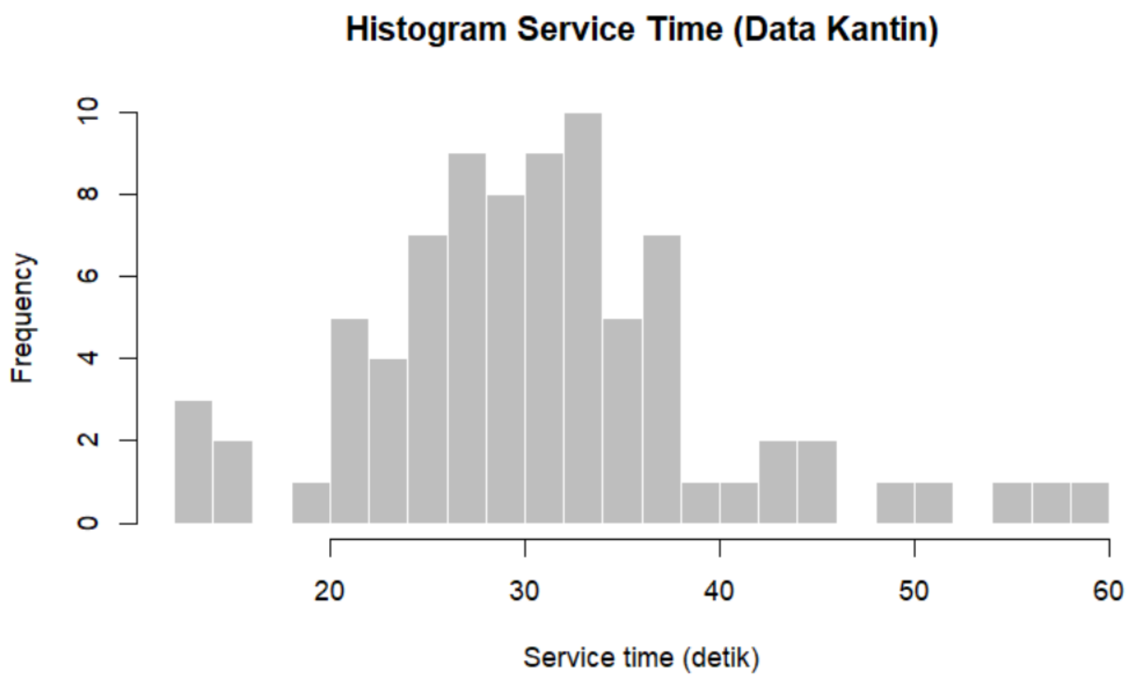


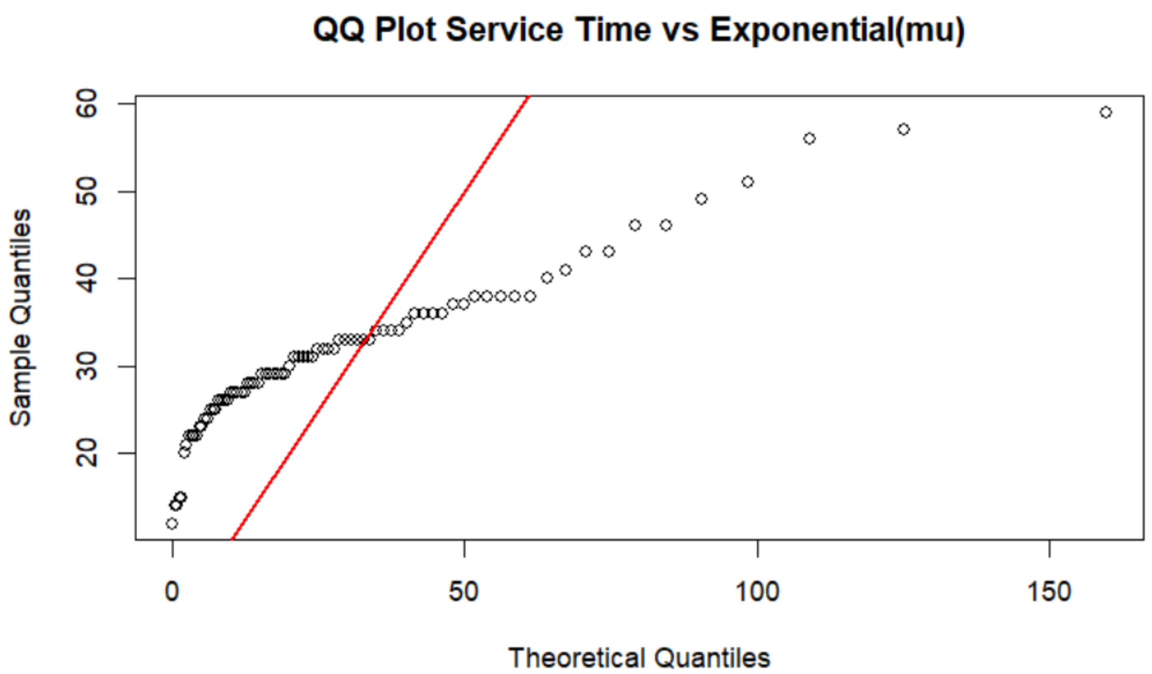


interarrival time berdistribusi eksponensial.

>Validasi 2: service time

```
{r}
# Histogram service time
hist(
  data$service_time,
  breaks = 30,
  main = "Histogram Service Time (Data Kantin)",
  xlab = "Service time (detik)",
  col = "gray",
  border = "white"
)

# QQ plot service time vs eksponensial(mu)
qqplot(
  qexp(ppoints(length(data$service_time)), rate = mu),
  data$service_time,
  main = "QQ Plot Service Time vs Exponential(mu)",
  xlab = "Theoretical Quantiles",
  ylab = "Sample Quantiles"
)
abline(0, 1, col = "red", lwd = 2)
```

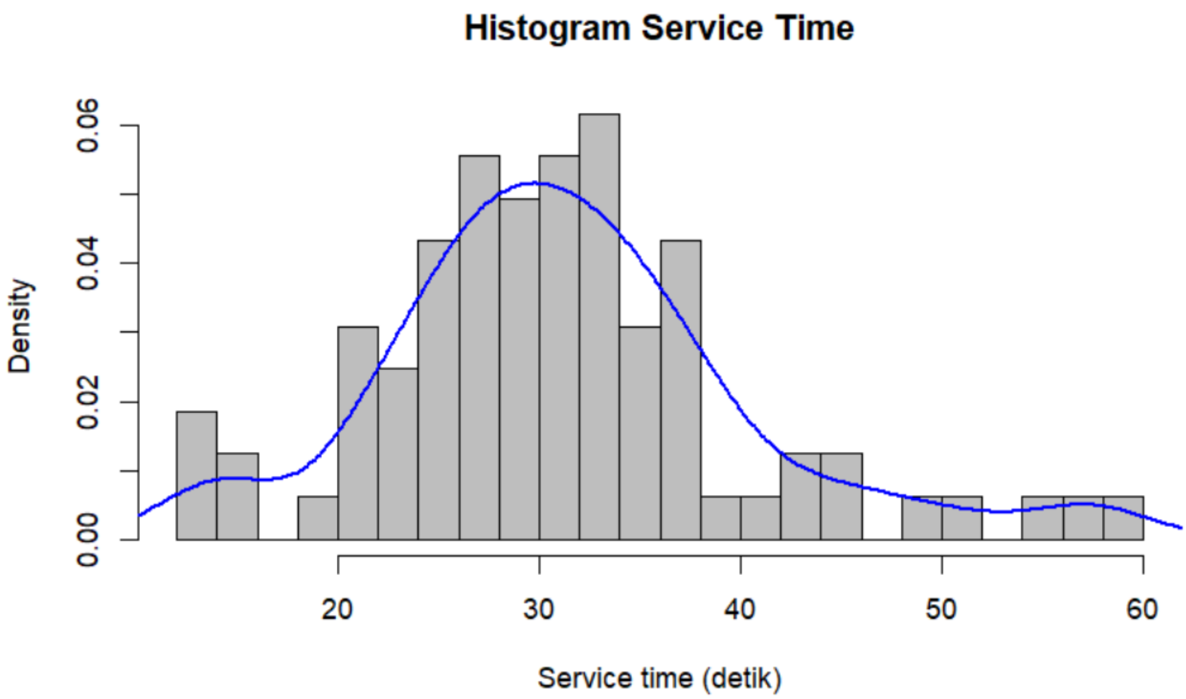




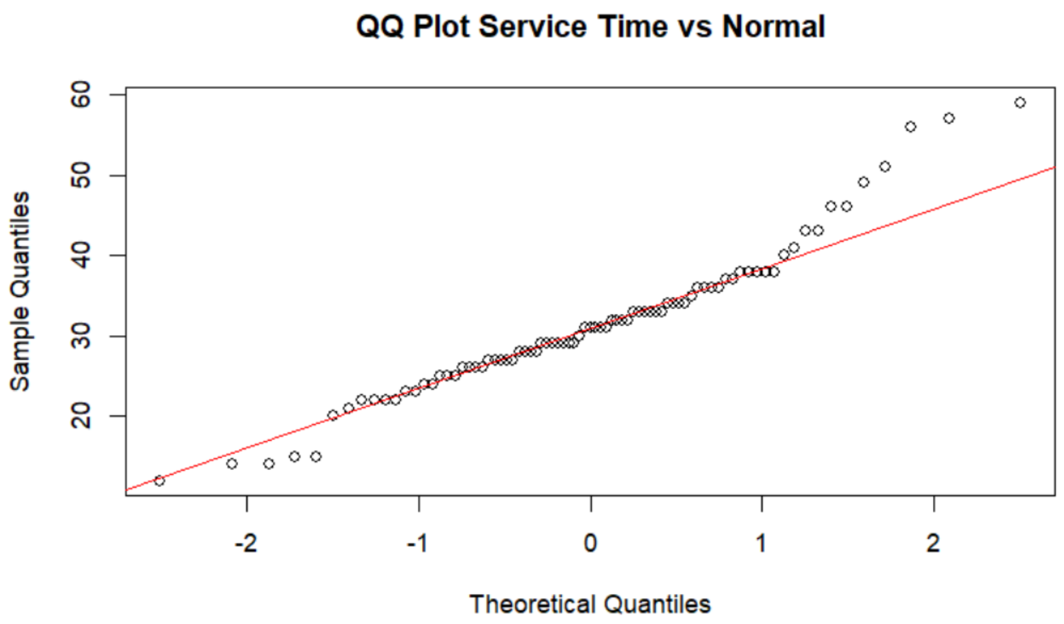
Service time tidak berdistribusi eksponensial, maka harus dicoba distribusi untuk menentukan model yang tepat.

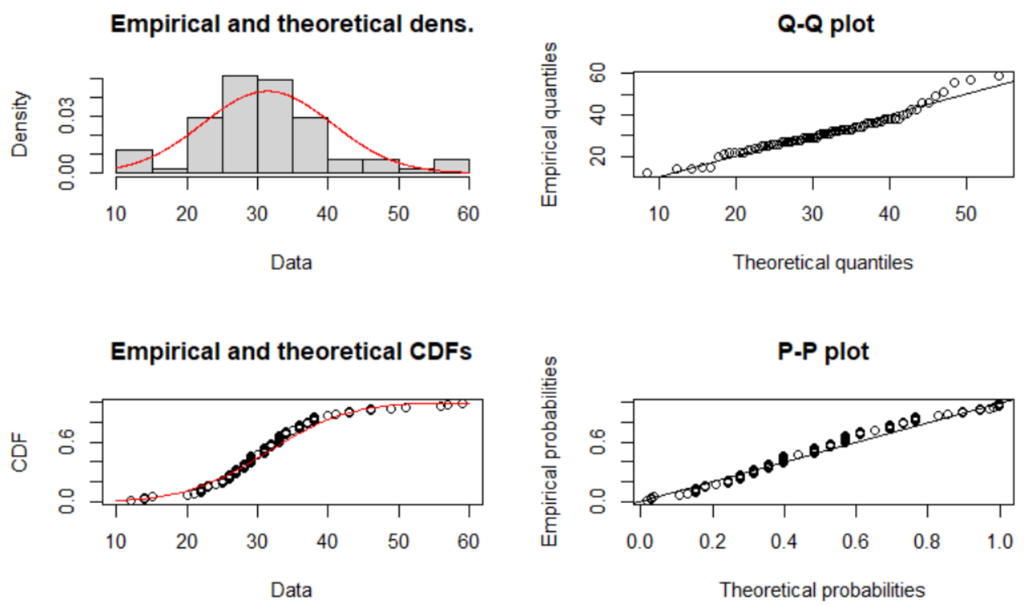
```
{r}
library(MASS)
library(fitdistrplus)

# data service time
x <- data$service_time
```

```
# 1. Histogram + density
hist(x, breaks = 20, col = "grey",
     main = "Histogram Service Time",
     xlab = "Service time (detik)",
     freq = FALSE)
lines(density(x), col = "blue", lwd = 2)
# 2. QQ plot normal
qqnorm(x, main = "QQ Plot Service Time vs Normal")
qqline(x, col = "red")
# 3. Uji normalitas (Shapiro-Wilk)
shapiro.test(x)
# 4. Fit distribusi (Normal, Lognormal, Gamma, Exponential)
fit_norm  <- fitdist(x, "norm")
fit_lnorm <- fitdist(x, "lnorm")
fit_gamma <- fitdist(x, "gamma")
# 5. Bandingkan dengan metrics AIC & BIC
fits <- list(norm = fit_norm,
             lognormal = fit_lnorm,
             gamma = fit_gamma)
sapply(fits, function(f) c(AIC = f$aic, BIC = f$bic))
# 6. Plot perbandingan fit
par(mfrow=c(2,2))
plot(fit_norm)
plot(fit_lnorm)
plot(fit_gamma)
```

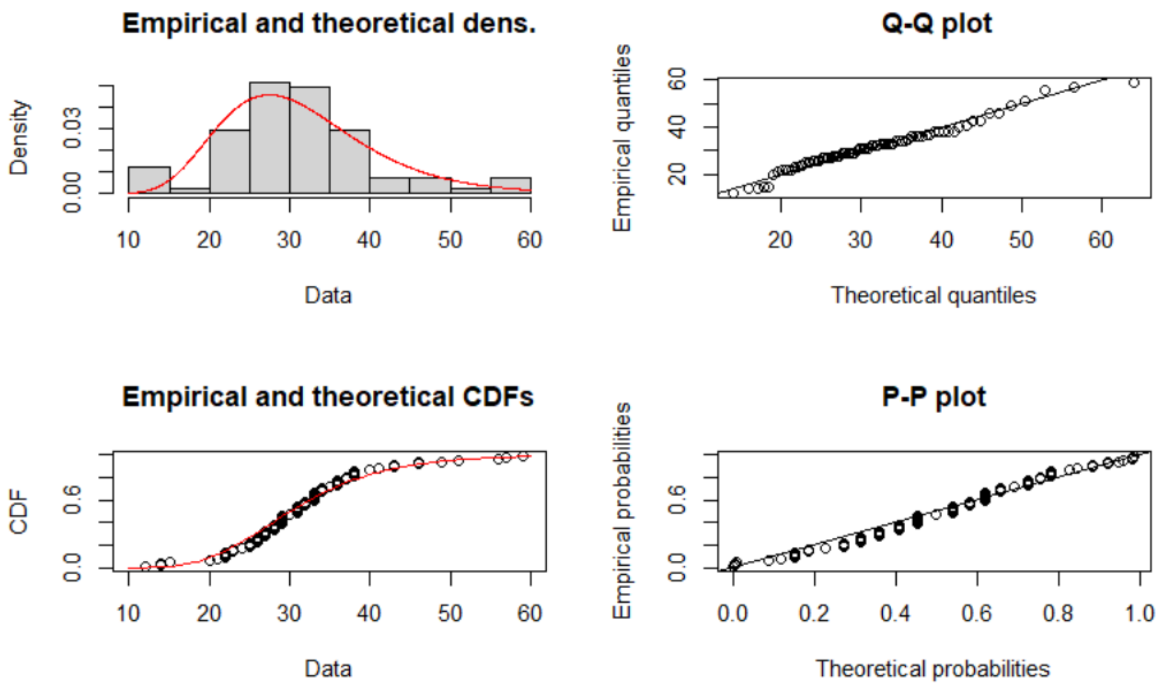


```
        Shapiro-Wilk normality test

data:  x
W = 0.95471, p-value = 0.006051

        norm lognormal    gamma
AIC 592.5658  591.2556 588.7036
BIC 597.3547  596.0445 593.4925
```

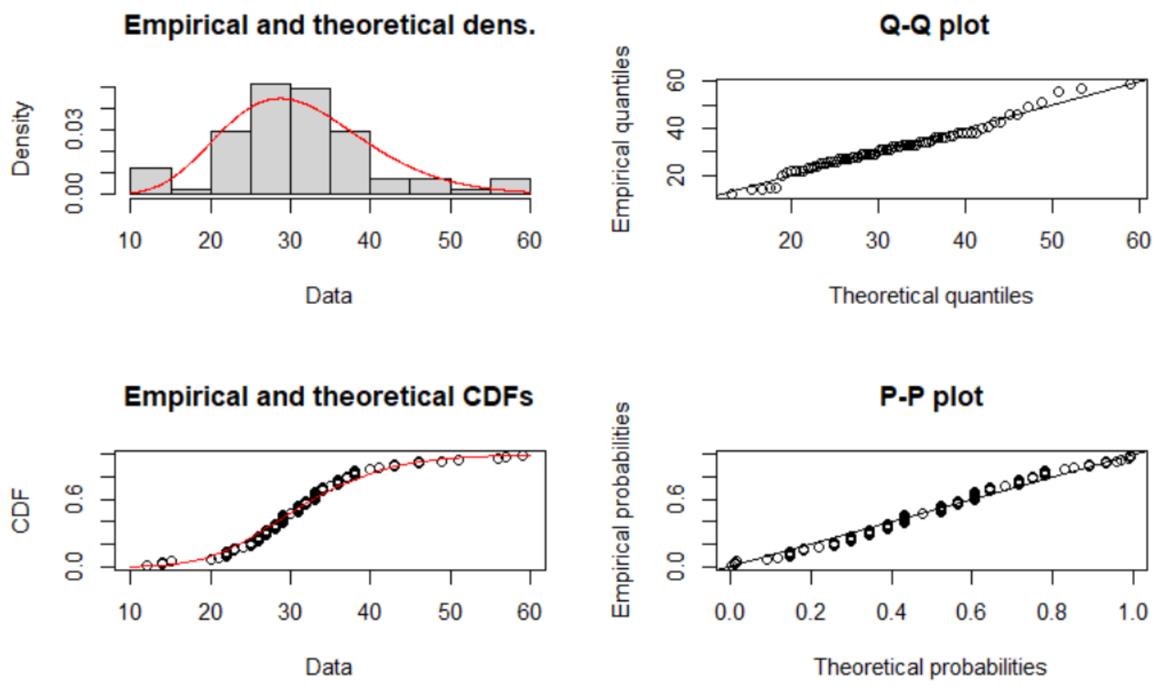




(Distribusi Normal)



(Distribusi LogNorm)



(Distribusi Gamma)

Distribusi service time terbukti **tidak normal**, terlihat dari uji Shapiro-Wilk yang menghasilkan **p-value = 0.006**, sehingga asumsi normalitas ditolak. Perbandingan model menggunakan nilai AIC dan BIC memperjelas arah pemilihannya. Distribusi Gamma memiliki nilai **AIC paling rendah (588.70)** dan **BIC paling rendah (593.49)**, lebih baik dibanding lognormal dan jauh lebih baik dibanding normal. Nilai AIC/BIC yang lebih kecil menandakan model yang lebih sesuai dengan pola data. Secara visual, kurva Gamma mengikuti bentuk histogram dengan paling konsisten, terutama di bagian ekor kanan yang lebih berat, sementara eksponensial dan normal gagal menangkap variasi ini. Berdasarkan bukti statistik dan visual tersebut, distribusi **Gamma** merupakan representasi paling akurat untuk service time pada data antrian kantin.

Hitung performa teori

```
{r}
# parameter hasil sebelumnya
lambda <- 0.02242781       # laju kedatangan
ES     <- 31.39506         # mean service time
VarS   <- 84.84198         # variansi service time
# hitung rho
rho <- lambda * ES
# hitung E[S^2]
ES2 <- VarS + ES^2
# hitung performa teori M/G/1
Lq <- (lambda^2 * ES2) / (2 * (1 - rho))
L  <- Lq + rho
Wq <- Lq / lambda
W  <- Wq + ES

Lq; L; Wq; W
```

```
[1] 0.9099448
[1] 1.614067
[1] 40.57216
[1] 71.96722
```

Dari parameter yang sudah dihitung sebelumnya. Didapatkan kombinasi yang menghasilkan tingkat utilisasi server **ρ ≈ 0.704**, yang berarti petugas aktif melayani sekitar 70 persen dari waktunya. Dengan memakai formula Pollaczek Khinchine untuk sistem M/G/1, diperoleh jumlah rata rata pelanggan dalam antrian (**Lq ≈ 0.910**) dan jumlah pelanggan rata rata dalam keseluruhan sistem (**L ≈ 1.614**). Waktu tunggu rata rata sebelum dilayani (**Wq ≈ 40.57 detik**) dan waktu total dalam sistem (**W ≈ 71.97 detik**) menggambarkan dampak variansi service time yang cukup besar. Secara keseluruhan, performa ini menunjukkan bahwa meskipun utilisasinya masih stabil, variabilitas pelayanan menyebabkan waktu tunggu yang cukup panjang bagi pelanggan.

Hitung performa empiris

```
{r}
# waktu tunggu (Wq) dan total waktu dalam sistem (W) dalam detik
data$Wq_emp <- as.numeric(difftime(data$wmd, data$wk, units = "secs"))
data$W_emp  <- as.numeric(difftime(data$wsd, data$wk, units = "secs"))
```

```
# hitung mean empirical
Wq_emp <- mean(data$Wq_emp, na.rm = TRUE)
W_emp  <- mean(data$W_emp,  na.rm = TRUE)

# dengan lambda yang sudah dihitung sebelumnya
lambda <- 0.02242781   # laju kedatangan /sec

# hitung Lq dan L empiris
Lq_emp <- lambda * Wq_emp
L_emp  <- lambda * W_emp

# hasil
Lq_emp; L_emp; Wq_emp; W_emp
```

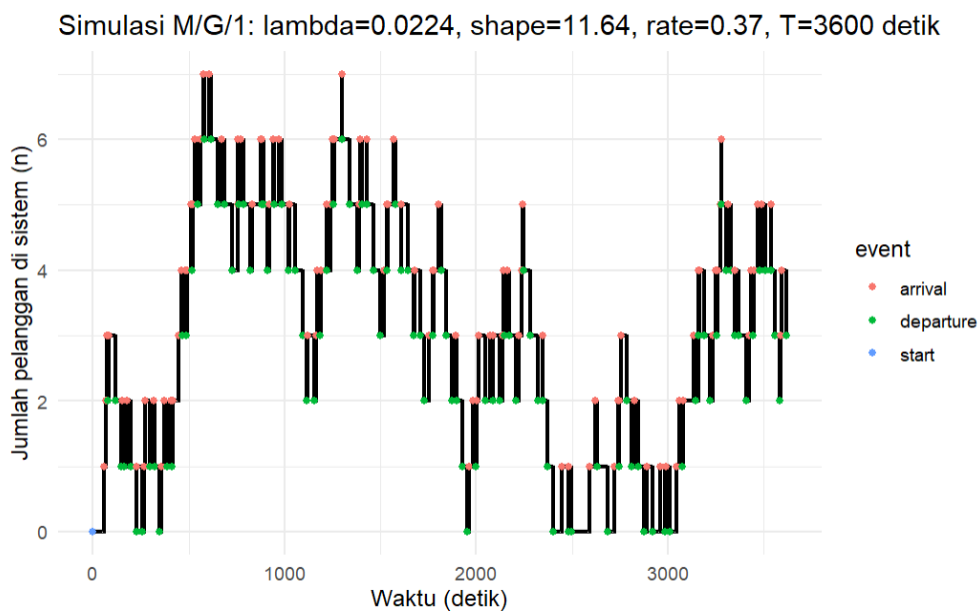
```
[1] 2.697429
[1] 3.401551
[1] 120.2716
[1] 151.6667
```

Hasil performa empiris yang dihitung langsung dari data aktual menunjukkan nilai yang jauh lebih besar nilai Lq_emp ≈ **2.70**, L_emp ≈ **3.40**, Wq_emp ≈ **120.27 detik**, dan W_emp ≈ **151.67 detik.** Sebaliknya, perhitungan performa teori menghasilkan estimasi Lq ≈ **0.91**, L ≈ **1.61**, Wq ≈ **40.57 detik**, dan W ≈ **71.97 detik**. Angka ini dihitung menggunakan parameter hasil estimasi, yaitu λ = **0.02243**, rata rata waktu pelayanan E[S] ≈ **31.40 detik**, serta variansinya Var(S) ≈ **84.84**. Perbedaan signifikan ini wajar muncul karena data riil mengandung variasi pelayanan yang tinggi dan fluktuasi kedatangan yang tidak selalu stabil. Variansi pelayanan yang besar membuat waktu tunggu aktual membengkak jauh lebih besar dari estimasi teori. Selain itu, data empiris hanya mewakili periode pengamatan tertentu, sehingga efek lonjakan kedatangan atau layanan lambat lebih terasa. Dengan kata lain, kedua hasil tetap valid: teori menggambarkan kondisi ideal dan stabil, sementara empiris mencerminkan dinamika nyata yang lebih bergejolak.

Simulasi M/G/1

```
{r}
library(tibble)
library(dplyr)
simulate_mg1 <- function(lambda, shape, rate, T = 3600){
  t <- 0
  n <- 0
  records <- tibble(time = 0, n = n, event = "start")
  next_arrival   <- rexp(1, rate = lambda)
  next_departure <- Inf

  while(t < T){
    if(next_arrival <= next_departure){
      # ARRIVAL
      t <- next_arrival
```

```
    n <- n + 1
    records <- add_row(records, time = t, n = n, event = "arrival")

    next_arrival <- t + rexp(1, rate = lambda)
    if(n == 1){
      # pelanggan ini langsung dilayani, service time ~ Gamma
      next_departure <- t + rgamma(1, shape = shape, rate = rate)
    }
  } else {
    # DEPARTURE
    t <- next_departure
    n <- max(0, n - 1)
    records <- add_row(records, time = t, n = n, event = "departure")
    if(n > 0){
      # masih ada yang ngantri, langsung generate service berikutnya
      next_departure <- t + rgamma(1, shape = shape, rate = rate)
    } else {
      next_departure <- Inf
    }
  }
  if(t > T) break
 }
 if(t < T){
  records <- add_row(records,
                     time  = T,
                     n     = records$n[nrow(records)],
                     event = "end")
 }
 records
}
shape_gamma <- fit_gamma$estimate["shape"]
rate_gamma  <- fit_gamma$estimate["rate"]
lambda      <- 0.02242781       # dari data, per detik
T_sim       <- 3600             # simulasi 1 jam
sim_mg1 <- simulate_mg1(lambda, shape_gamma, rate_gamma, T = T_sim)
# plot jumlah pelanggan di sistem terhadap waktu
library(ggplot2)
sim_mg1_plot <- sim_mg1 %>%
 arrange(time) %>%
 mutate(next_time = lead(time, default = T_sim),
        duration  = next_time - time)
ggplot(sim_mg1, aes(x = time, y = n)) +
 geom_step(linewidth = 1) +
 geom_point(aes(col = event), size = 1.5) +
 labs(
  title = sprintf("Simulasi M/G/1: lambda=%.4f, shape=%.2f, rate=%.2f, T=%d detik",
                  lambda, shape_gamma, rate_gamma, T_sim),
  x = "Waktu (detik)",
  y = "Jumlah pelanggan di sistem (n)"
 ) +
 theme_minimal(base_size = 12)
```



Simulasi dilakukan untuk menguji apakah model antrian yang dibangun mampu mereproduksi pola kinerja yang mendekati kondisi nyata. Dengan menggunakan parameter kedatangan dan distribusi waktu pelayanan yang diestimasi dari data, simulasi menghasilkan nilai Lq, L, Wq, dan W yang tidak persis sama dengan data empiris, tetapi menunjukkan tren yang konsisten. Hal ini menunjukkan bahwa model mampu menangkap karakter utama dari dinamika antrian, meskipun masih terdapat selisih akibat keterbatasan durasi simulasi dan variasi acak pada setiap replikasi. Secara praktis, simulasi ini berfungsi sebagai jembatan antara perhitungan teori dan hasil empiris, sehingga membantu memvalidasi bahwa model antrian yang digunakan masih relevan untuk menggambarkan sistem kantin.

Hitung performa simulasi M/G/1

```
{r}
sim <- sim_mg1   # nama data hasil simulasi lo
# waktu tunggu dan waktu sistem berbasis simulasi
sim <- sim %>%
  arrange(time) %>%
  mutate(
    next_time = lead(time),
    dt        = next_time - time
  )

# estimasi L (jumlah rata rata pelanggan dalam sistem)
L_sim <- sum(sim$n * sim$dt, na.rm = TRUE) / max(sim$time)
# Lq (yang menunggu, bukan yang sedang dilayani)
Lq_sim <- sum(pmax(sim$n - 1, 0) * sim$dt, na.rm = TRUE) / max(sim$time)
# W & Wq dari Little's Law
lambda_sim <- lambda   # sama kayak laju kedatangan asli
W_sim  <- L_sim  / lambda_sim
Wq_sim <- Lq_sim / lambda_sim
```

```
L_sim; Lq_sim; W_sim; Wq_sim
```

```
[1] 2.994976
[1] 2.09527
[1] 133.5385
[1] 93.42284
```

Perbandingan Performa Teori, Empiris, dan Simulasi

```
{r}
comparison <- tibble(
  Metrik = c("Lq", "L", "Wq (detik)", "W (detik)"),
  Teori = c(Lq, L, Wq, W),
  Empiris = c(Lq_emp, L_emp, Wq_emp, W_emp),
  Simulasi = c(Lq_sim, L_sim, Wq_sim, W_sim)
)
comparison
```

A tibble: 4 × 4

Metrik <chr>	Teori <dbl>	Empiris <dbl>	Simulasi <dbl>
Lq	0.9099448	2.697429	2.095270
L	1.6140672	3.401551	2.994976
Wq (detik)	40.5721646	120.271605	93.422839
W (detik)	71.9672246	151.666667	133.538480

4 rows

Berdasarkan Tabel perbandingan performa, terlihat bahwa nilai rata-rata jumlah pelanggan dalam antrian (Lq), jumlah pelanggan dalam sistem (L), waktu tunggu antrian (Wq), dan waktu total di sistem (W) yang diperoleh dari data empiris secara konsisten lebih besar dibandingkan hasil perhitungan teori. Hal ini sejalan dengan karakteristik data pelayanan yang memiliki variansi tinggi, sehingga sistem nyata menjadi lebih sensitif terhadap fluktuasi kedatangan dan ketidakkonsistenan waktu pelayanan. Hasil simulasi berada pada kisaran nilai yang sejalan dengan pola tersebut dan memberikan gambaran tambahan mengenai perilaku sistem ketika diasumsikan mengikuti proses kedatangan dan pelayanan yang telah diperkirakan dari data. Dengan demikian, teori memberikan batas kinerja ideal, simulasi merepresentasikan perilaku sistem di bawah model antrian yang diasumsikan, sedangkan perhitungan empiris menggambarkan kondisi operasional yang benar-benar terjadi di kantin.

Kesimpulan

Hasil empiris menunjukkan bahwa waktu tunggu pelanggan relatif tinggi pada periode makan siang, yaitu waktu pengambilan data dalam penelitian ini. Fokus analisis memang diarahkan pada satu kasir saja meskipun kantin memiliki dua kasir aktif. Pendekatan ini tetap relevan karena pada jam makan siang kedua kasir menghadapi lonjakan pelanggan yang sama, sehingga satu kasir sudah cukup mewakili pola antrean yang terjadi.

Beberapa hal seperti waktu pelayanan dapat dipercepat bila pelanggan menyiapkan uang pas terlebih dahulu atau memastikan metode pembayaran digital berjalan lancar, misalnya dengan memastikan jaringan stabil sebelum tiba giliran. Langkah sederhana seperti ini dapat menurunkan variasi waktu pelayanan, yang menurut hasil perhitungan menjadi salah satu faktor utama memanjangnya antrean.

Dari sisi operasional, memastikan perangkat pembayaran tetap responsif juga membantu mengurangi fluktuasi waktu pelayanan. Mengingat tingginya intensitas kedatangan khusus pada jam makan siang, pengurangan variabilitas pelayanan bisa berdampak lebih signifikan dibanding menambah kasir.