

Recitation 8

Forecasting Arrivals (demand for service)

Problem Setup

Days are divided into time intervals, with an assumed constant arrival rate over an interval.

Practice: 15 min, 30 min, 1 hour.

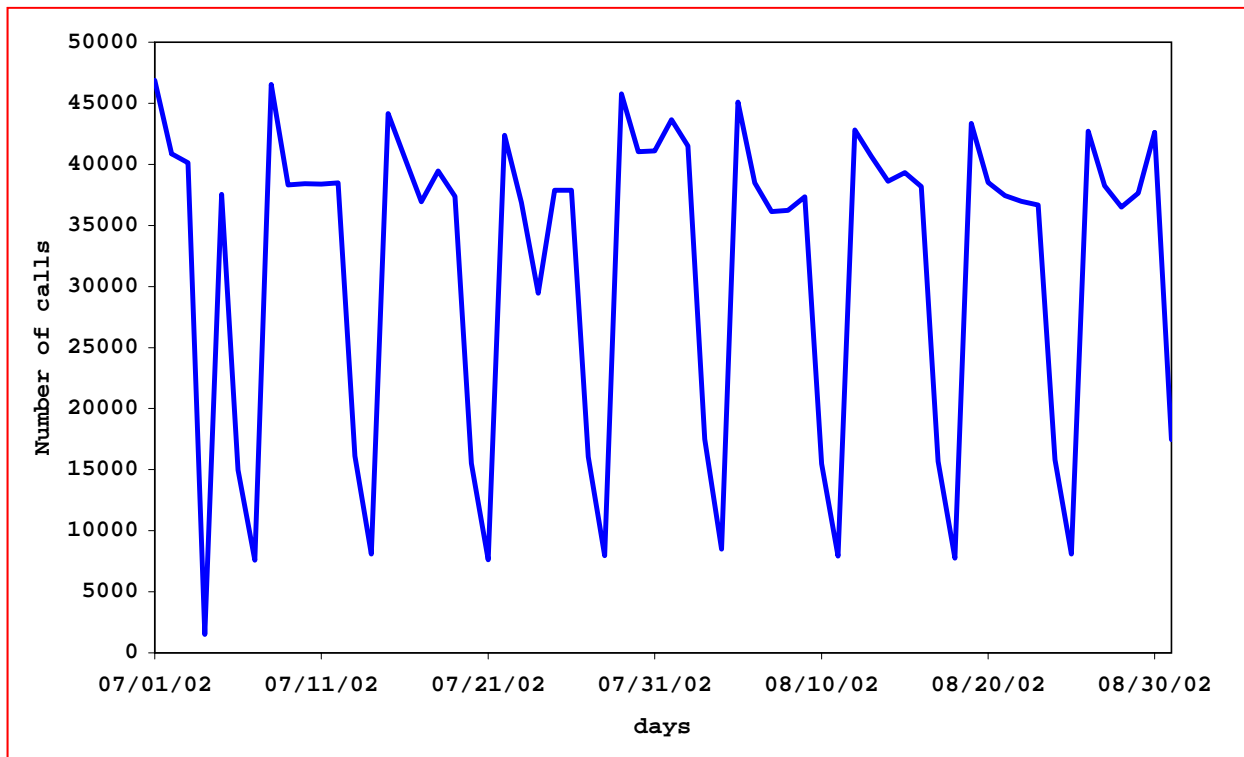
N_{jk} = # of arrivals, during time-interval k , on day j .
Assume J days overall, with K intervals per day.

- **One-day-ahead** prediction:
 $N_{1,\cdot}, \dots, N_{j-1,\cdot}$ known. Predict N_{j1}, \dots, N_{jK} .
- **Several days (weeks) ahead** prediction.
- **Within-day** prediction.
 $N_{1,\cdot}, \dots, N_{j-1,\cdot}, N_{j1}, \dots, N_{j,k-1}$ known.
Predict N_{jk}, \dots, N_{jK} .

Practice: Do all the above, via nested rolling horizon (Weekly, Daily, Hourly).

Daily Call Volumes

Daily Call Volumes at a U.S. Bank. July-August 2002.

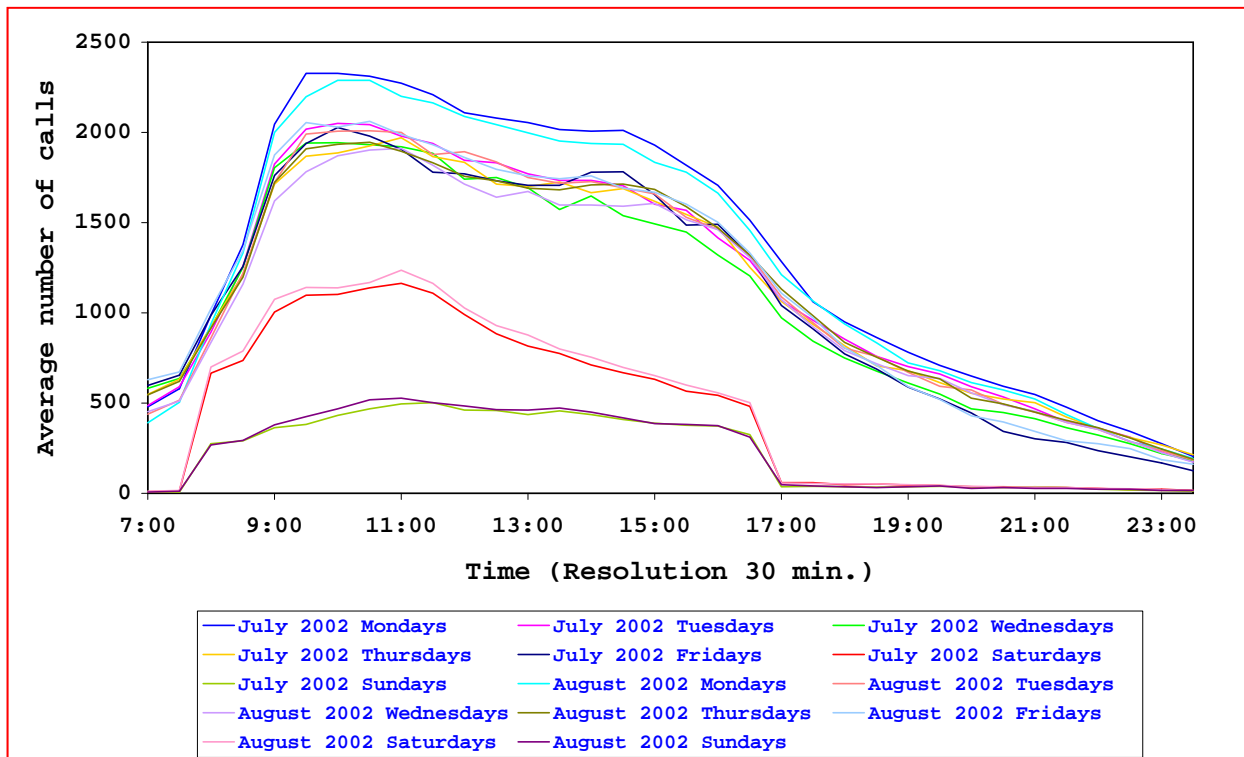


Guiding Questions.

- Do we observe trends?
- Do we observe differences between days-of-week?
- Do we observe other periodic effects?
- Do we observe any “unusual” days?

Intra-day Call Volumes

Intraday Call Volumes at a U.S. Bank. July-August 2002.



Guiding Questions.

- Noticeable pattern? (eg. peak time, closing).
- Do days-of-week differ in shape? volume?
- Are volume/shape stable over time?

Forecasting: Simple Methods

Most recent observation.

F_{jk} = most recent “similar” call volume.

Example: $F_{jk} = N_{j-7,k}$ (previous week).

Moving average.

Average of several (not too many) recent “similar” call volumes.

Most Recent, plus Yesterday’s Correction.

Example: Factor accounting for a “busy yesterday”.

What about sophisticated forecasting methods?

Active research.

In the **Homework**, we shall compare the performance of **simple methods against** (given results of) **sophisticated methods**.

Forecasting: Goodness-of-Fit

How to compare different methods?

N_{jk} – number of calls (day j , interval k);

F_{jk} – forecast.

Two ways to quantify forecasting accuracy:

1. Root Mean-Square Error (RMSE)

For each day j , calculate:

$$RMSE_j = \sqrt{\frac{1}{K} \sum_{k=1}^K (N_{jk} - F_{jk})^2}.$$

$$RMSE = \sum_{j=1}^J \frac{RMSE_j}{J}.$$

2. Average Percent Error (APE)

$$APE_j = \frac{100}{K} \cdot \sum_{k=1}^K \frac{|N_{jk} - F_{jk}|}{N_{jk}}.$$

$$APE = \sum_{j=1}^J \frac{APE_j}{J}.$$

Time Series Models

Standard mathematical technique for forecasting.

A course on Time Series is taught at the faculty of IE&M.

Examples of Time Series Models

AR(1) (Auto-Regressive) model:

$$N_t = c + \phi N_{t-1} + Z_t, \quad Z_t \sim N(0, \sigma^2) \text{ iid}, \quad |\phi| < 1.$$

MA(1) (Moving Average) model:

$$N_t = \mu + \theta Z_{t-1} + Z_t, \quad Z_t \sim N(0, \sigma^2) \text{ iid.}$$

ARMA (Auto-Regressive Moving Average) models. Combine both features.

ARIMA (Auto-Regressive Integrated Moving Average) models that treat “trend” and “seasonable components”.

Variance Stabilizing Transformation

In time series model the variance of random error is assumed stable. In fact, it depends on the arrival volume.

But:

If $N \sim \text{Pois}(\lambda)$ then

$$Y = \sqrt{N + \frac{1}{4}} \approx \text{Normal}(\sqrt{\lambda}, 1/4),$$

for large λ .

Arrival Rate to Israeli Telecom: Example of Time Series Model

M.Sc. of Sivan Aldor (2005).

Used to calculate forecasts in the Homework.

$$Y_{dk} = \alpha_{q(d)} + \pi_{q(d)k} + \text{Billing}_{y(d)} + \tau_d + \beta_k + \epsilon_{dk},$$

where

Y_{dk} – arrival volume on day d during interval k
(after square-root transformation);

$q(d)$ – day of week;

$\alpha_{q(d)}$ – day-of-week effect (fixed);

$\pi_{q(d)k}$ – time-interval effect (fixed);

$\text{Billing}_{y(d)}$ – connected with monthly billing cycles (fixed);

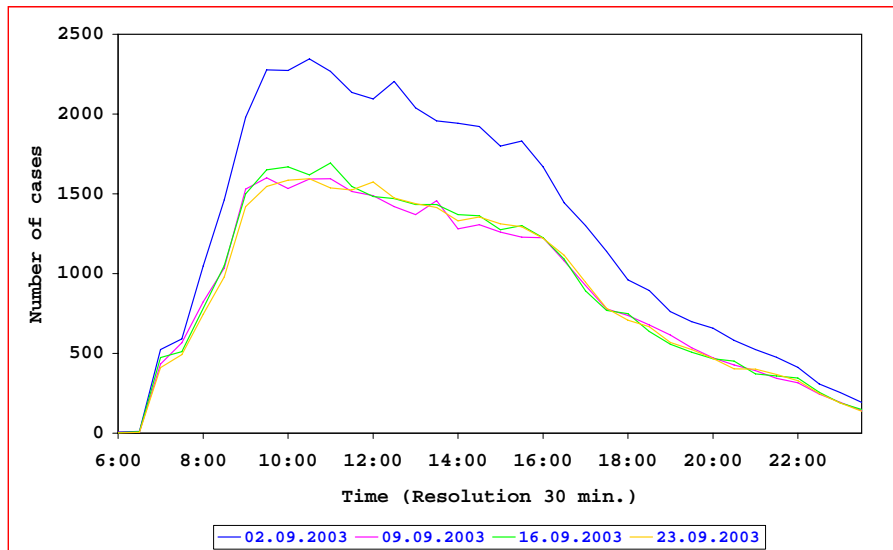
τ_d – daily random effect, based on AR(1) model;

β_k – interval random effect, based on AR(1) model;

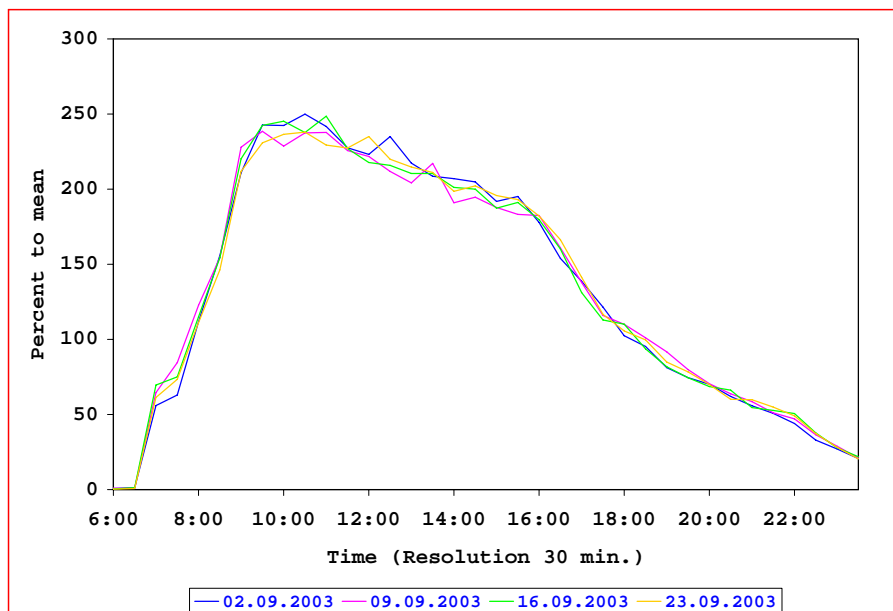
ϵ_{dk} – random error.

Stability of Intraday Arrival Shape

US Bank: Retail, September 2003 Daily Arrival Rate on Tuesdays



Percent to Mean



Stability of Intraday Arrival Shape: Application to Forecasting

Formal Setup: Consider days with “similar shape” (Sundays, Fridays, weekdays etc).

Assume that arrival process on these days is time-inhomogeneous Poisson with arrival rate given by $C\lambda_{\%}(t)$, $0 \leq t \leq 24$, where $\int_0^{24} \lambda_{\%}(t)dt = 1$.

C = daily arrival volume – random variable.

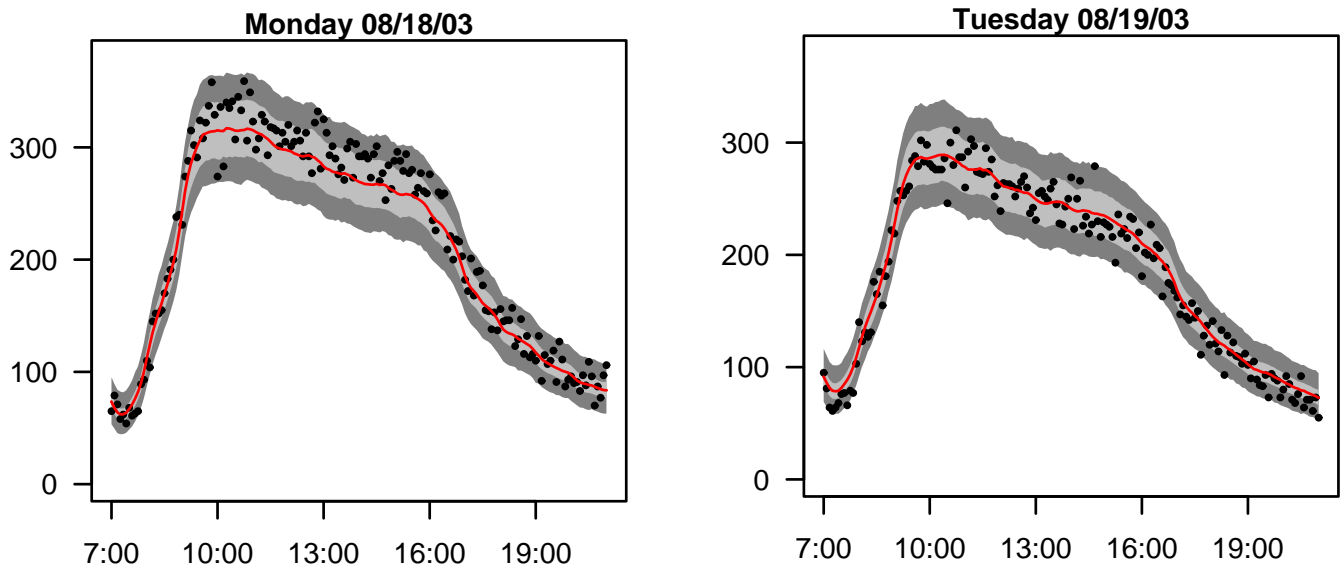
$\lambda_{\%}(t)$ = intraday arrival shape – deterministic.

Why important in forecasting: allows to separate prediction of daily arrival volume (**time series**) and estimation of arrival shape (**historical averages**).

You will use a simple method of this type in the Homework.

Forecast Performance: Example

US Bank: Forecast Performance (Weinberg, Brown, Stroud, 2005)



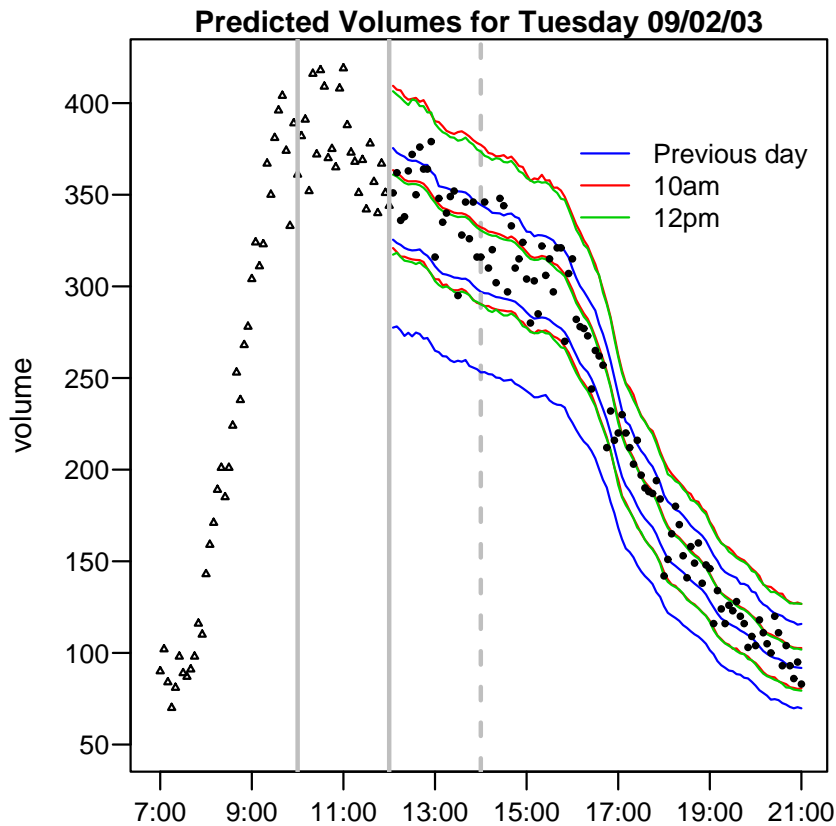
Wider confidence intervals for **number of calls**.

Narrower confidence intervals for **arrival rate**
(Poisson parameter).

Note: **staffing** models require an **arrival rate** as input.

Within-Day Updating

Comparison between Day-Ahead and Within-Day Predictions (Weinberg, Brown, Stroud, 2005)



Conclusion: Morning information is important but no significant difference between 10am and 12am.

Concluding Remarks

- Best models can be very sophisticated.
- Yet simple models are very useful.
- How to account for trends and periodic effects.
- Different types of periodic effects are observed.
- Call volume at “previous day” is important.
- Stability of intraday arrival shape can be helpful.
- Morning information is important.