Forecasting Arrivals (demand for service)

- Forecasting Arrivals
  - Problem Setup and Goodness-of-Fit (p.2-5)
  - Simple Methods (p.6)
  - Time Series Models (p.7-8)
  - EX. Arrival Rate to Israeli Telecom (p. 9-14)

Forecasting Arrivals

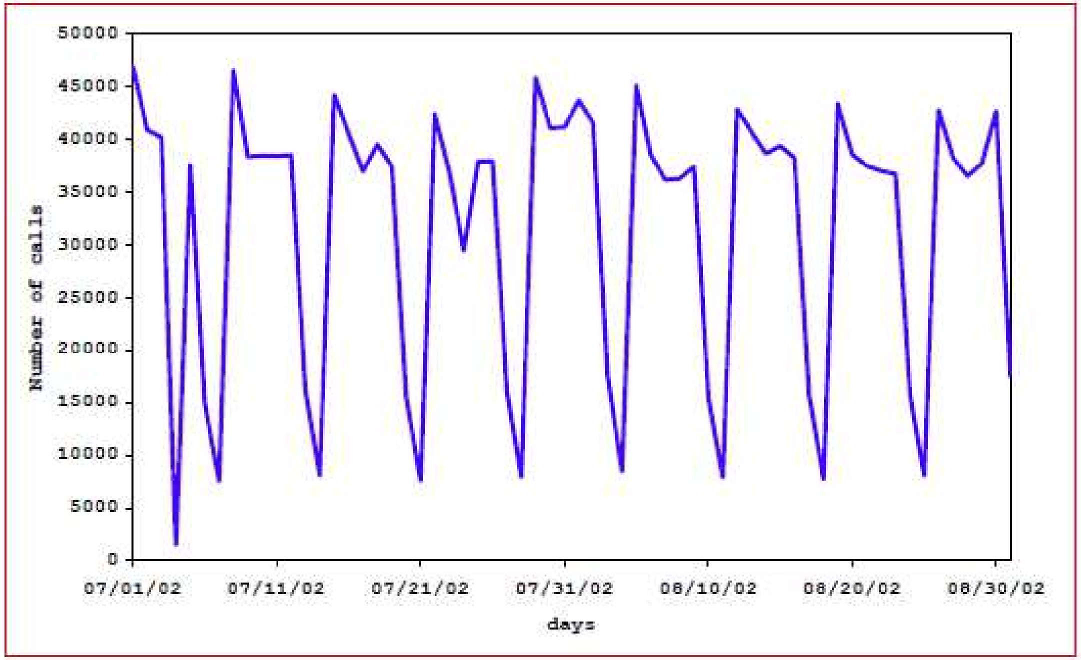
- Days are divided into time intervals (e.g., 15 min, 30 min, 1 hour), in which we assume a constant arrival rate

- N_{jk} = # of arrivals, during time-interval k, on day j.
  - Assume J days overall, with K intervals per day
  - Examples:
    - One-day-ahead prediction: $N_1, \dots, N_{j-1}$ known. Predict $N_{j1}, \dots, N_{jK}$
    - Several days (weeks) ahead prediction.
    - Within-day prediction:
      - $N_1, \dots, N_{j-1}, N_{j1}, \dots, N_{j,k-1}$ known. Predict $N_{jk}, \dots, N_{jK}$

- We can do all the above, via nested rolling horizon
  - Weekly, Daily, and Hourly

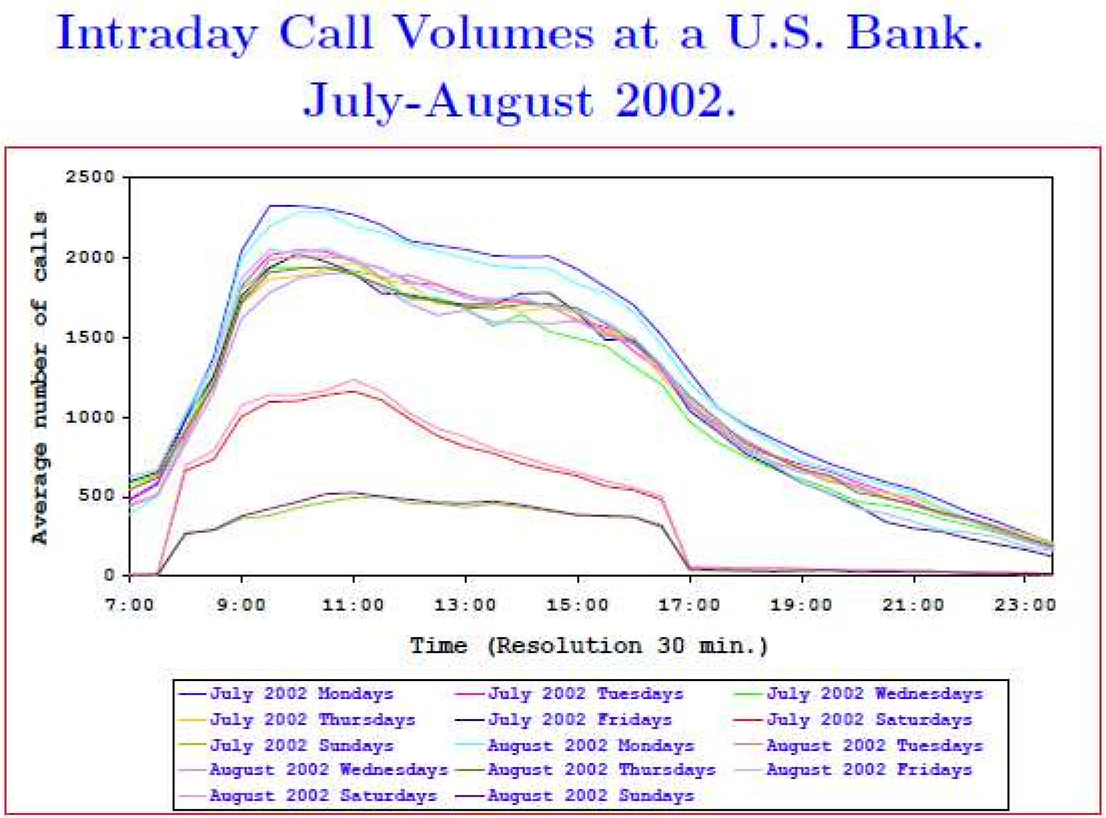
Forecasting Arrivals (2)

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?

Forecasting Arrivals (3)



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

Forecasting Arrivals (4)

- 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)

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

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

Simple Methods

Method 1. Last observation. Let the forecast be equal to the last call volume of the same type, meaning the call volume during the same time interval of the same day-of-week on the previous week: $F_{jk} = N_{j-6,k}$.

Method 2. Moving average. Let the forecast be equal to the average of five last call volumes of the same type (previous 5 weeks).

$$e.g., F_{jk} = \frac{N_{j-7,k} + N_{j-14,k} + \ \dots + N_{j-35,k}}{5}$$

Method 3. Using yesterday's call volume during the considered time interval. Implement forecasting formula:

$$F_{jk} = N_{j-1,k} + (F_{2,jk} - F_{2,j-1,k})$$

Where $F_{2,jk}$ and $F_{2,j-1,k}$ are forecasts by Method 2. We add the expression in brackets in order to incorporate difference between days-of-week in our forecast.

Time Series Models

- Standard mathematical technique for forecasting

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".

Time Series Models (2)

- In time series models, the variance of random error is assumed to be stable.

- If $\mathbf{N \sim Poisson\ (\lambda)}$, $\mathrm{Var(N)} = \lambda$ (i.e., the variance varies with the mean).
  - Variance Stabilizing Transformation
    - find a simple function to create new values $\mathrm{y = f(x)}$ such that the values y is not related to their mean value for Poisson Process
    - If $\mathrm{N{\sim}Poisson(\lambda)}$, using $\sqrt{\mathrm{N + c}}$ improves the variance stabilizing properties [Anscombe, *Biometrika*, 1948. - Uses Taylor series expansion and analyzes the remainder term]
  - In addition, we want to choose the constant c so that the mean of $\sqrt{\mathrm{N + c}}$ is 'closest' to $\sqrt{\lambda}$ (mean matching)
    - $\mathbf{Y} = \sqrt{\mathbf{N} + \frac{\mathbf{1}}{\mathbf{4}}} \approx \mathbf{Normal}\ (\sqrt{\boldsymbol{\lambda}}, \frac{\mathbf{1}}{\mathbf{4}})$ for large λ (approximately $\lambda \geq 10$)

    [Brown et al., *Probab. Theory Relat. Fields*, 2010.]

EX. Arrival Rate to Israeli Telecom

- METHOD 1 [Aldor-Noiman et al., *The Annals of Applied Statistics,* 2009.]:

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

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;

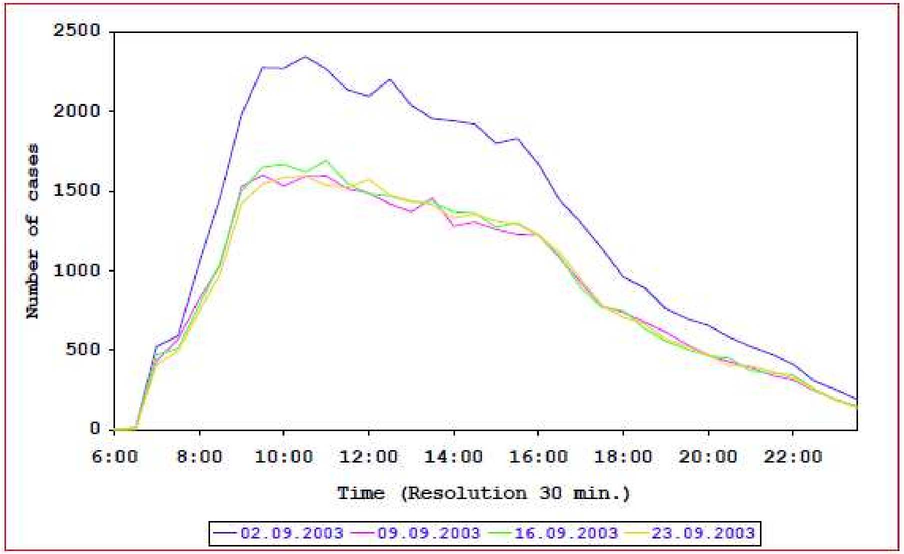
β_k – interval random effect, based on AR(1) model;

ϵ_{dk} – random error.

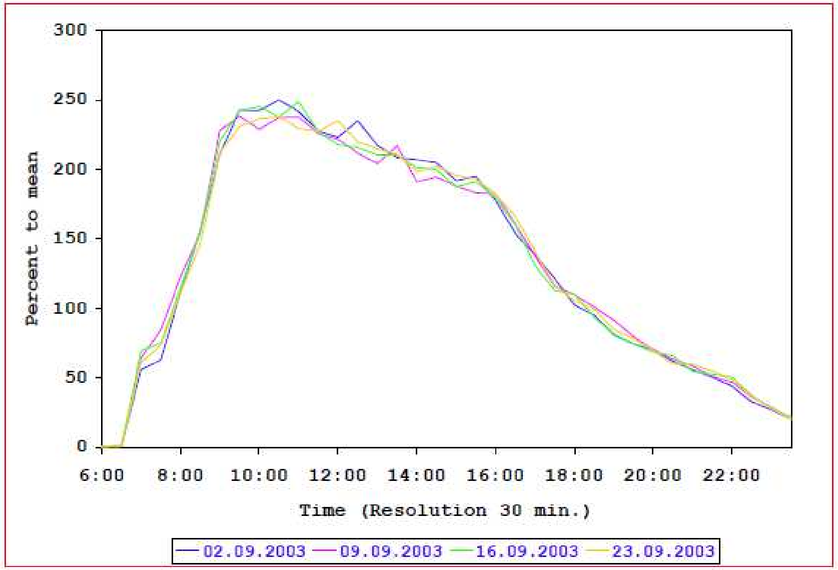
EX. Arrival Rate to Israeli Telecom (2)

- METHOD 2: Use Intraday Arrival Shape
  - US Bank: Retail, September 2003

Daily Arrival Rate on Tuesdays



Percent to Mean



EX. Arrival Rate to Israeli Telecom (3)

- Intraday Arrival **Shape** – application to forecasting

    - 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 $C\lambda_{\%}(t),\ 0 \le t \le 24$, where $\int_0^{24} \lambda_{\%}(t)dt = 1$.

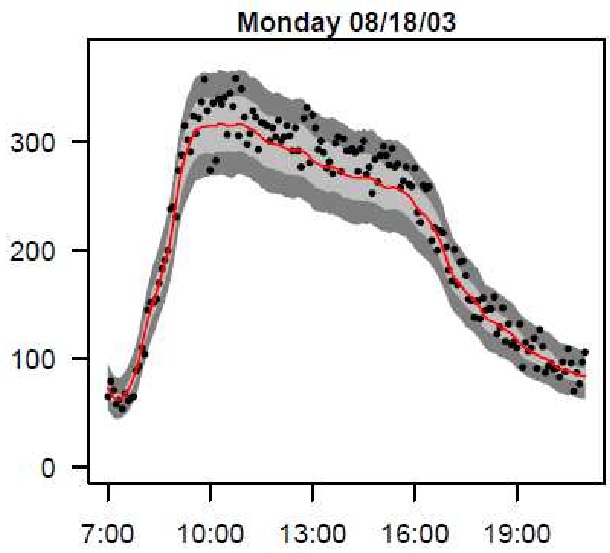
$\boldsymbol{C}$ **= daily arrival volume** – random variable.

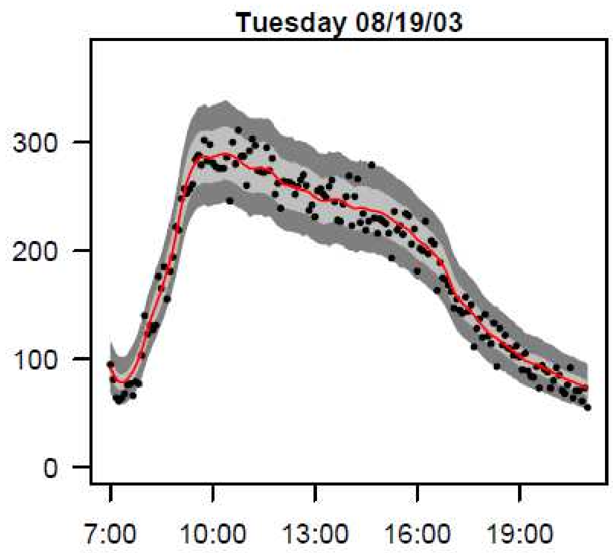
$\boldsymbol{\lambda_{\%}(t)}$ **= intraday arrival shape** – deterministic.

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

EX. Arrival Rate to Israeli Telecom (4)

- Forecast Performance (Weinberg, Brown, Stroud 2005)





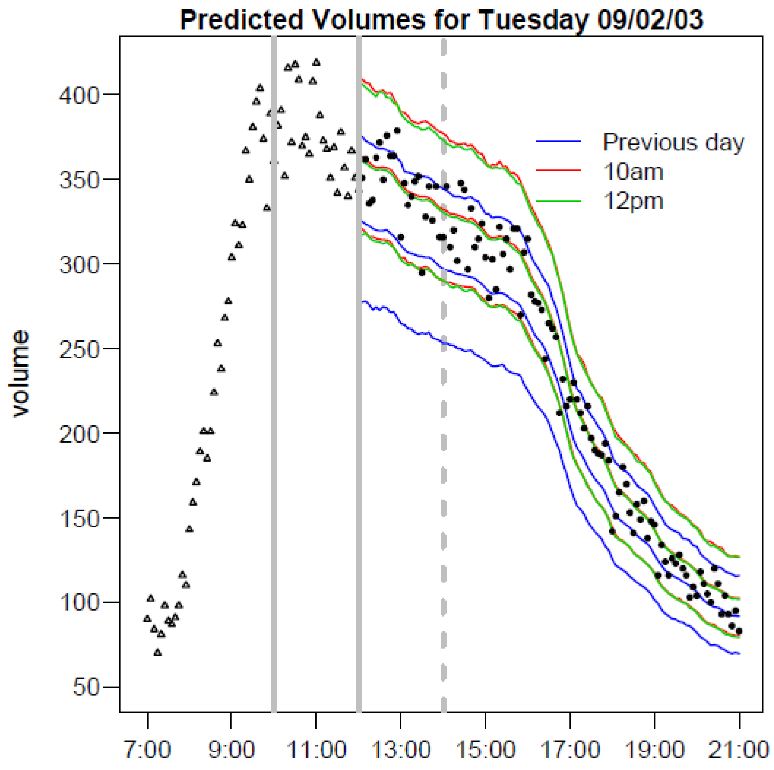
Wider confidence intervals for **number of calls**.

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

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

EX. Arrival Rate to Israeli Telecom (5)

- Within-day update
  - Comparison between day-ahead and within-day predictions



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

EX. Arrival Rate to Israeli Telecom (5)

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.