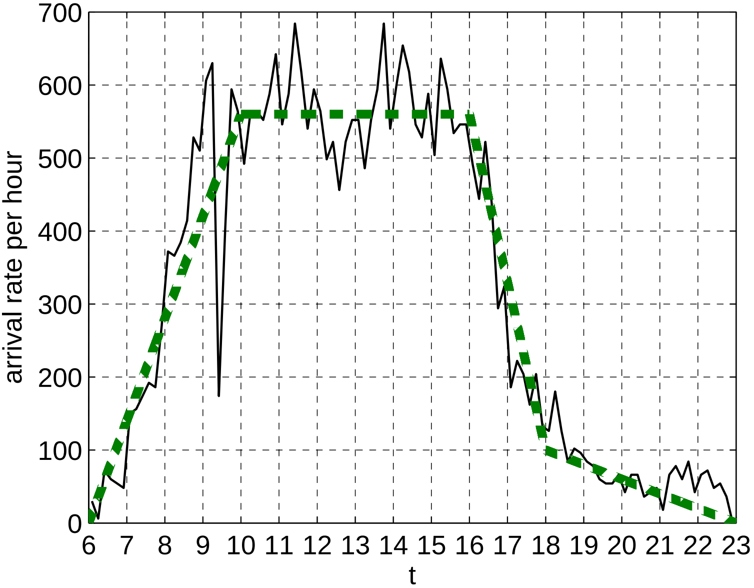
Fitting Queueing Models to Service System Data: What Arrival Process Model is Appropriate?

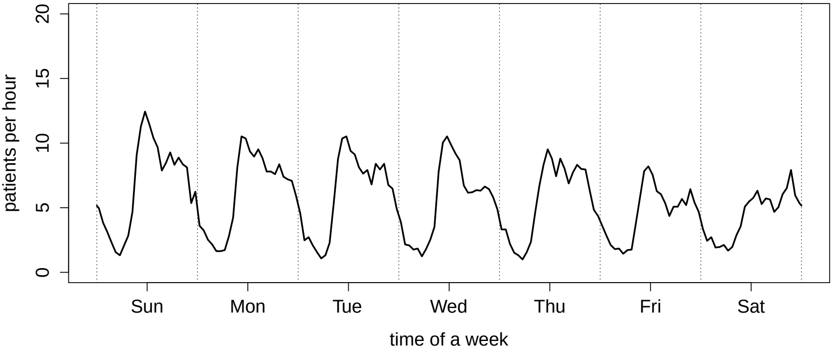
Ward Whitt
Department of Industrial Engineering and Operations Research
Columbia University, New York, NY 10027
URL: http://www.columbia.edu/∼ww2040

joint work with
Song-Hee Kim, USC Marshall School of Business
(former doctoral student)

Call Arrival Rate to a U.S. Bank Service Center over a Day



Arrival Rate to an Israeli ED over a Week



What Arrival Process Model is Appropriate

- for the bank?

What Arrival Process Model is Appropriate

- for the bank?
- hopefully a homogeneous Poisson process (**PP**)
  (Treating as stationary over half hours because service times are short, i.e., using the Pointwise-Stationary Approximation)

What Arrival Process Model is Appropriate

- for the bank?
- hopefully a homogeneous Poisson process (**PP**)
  (Treating as stationary over half hours because service times are short, i.e., using the Pointwise-Stationary Approximation)
- and for the ED?

What Arrival Process Model is Appropriate

- for the bank?
- hopefully a homogeneous Poisson process (**PP**)
  (Treating as stationary over half hours because service times are short, i.e., using the Pointwise-Stationary Approximation)

- and for the ED?
- hopefully a nonhomogeneous Poisson process (**NHPP**)
  (Patient Lengths of Stay (LoS) are too long for PSA.)

What Arrival Process Model is Appropriate

- for the bank?
- hopefully a homogeneous Poisson process (**PP**)
  (Treating as stationary over half hours because service times are short, i.e., using the Pointwise-Stationary Approximation)

- and for the ED?
- hopefully a nonhomogeneous Poisson process (**NHPP**)
  (Patient Lengths of Stay (LoS) are too long for PSA.)

- **But are these Poisson models appropriate?**

A Nonhomogeneous Poisson Process (NHPP)

- **no batches**: Arrival occur one at a time.
- **independent increments**: For all $k \geq 2$ and $0 \leq t_1 \leq t_2 \leq t_3 \leq \cdots t_{2k}$, $A(t_2) - A(t_1)$ (number of arrivals in $[t_1, t_2]$), $A(t_4) - A(t_3), \ldots, A(t_{2k}) - A(t_{2k-1})$ are k independent random variables.
- **increments have a Poisson distribution**: $P(A(t_2) - A(t_1) = k) = \frac{e^{-m(t_1,t_2))} m(t_1,t_2)^k}{k!}$
- **mean is the integral of the arrival rate $\lambda(t)$**: $m(t_1, t_2) = \int_{t_1}^{t_2} \lambda(s)\, ds,\ t \geq 0$.
- The **Poisson process** (PP) is the special case in which: $\lambda(t) = \lambda$ (constant). Unlike the NHPP, the PP has **stationary increments**.

Justification for Poisson: People Acting Independently

Theorem

(The Law of Rare Events) The superposition of n independent i.i.d. nonstationary (stationary) point processes with intensities $\lambda(t)/n$ (λ/n) converges in distribution to an NHPP (PP) with arrival rate function $\lambda(t)$ (λ) as $n \rightarrow \infty$.

- Often called the Palm-Khintchine theorem.
- The component processes need not be renewal, as often assumed.
- The component processes need not be identical.

Examples Where PP and NHPP Are Suspect

- Likely to be smoother, **less variable or less bursty**:
  - **scheduled arrivals**, as at doctor's office.
  - **enforced separation**, as in landings at airports.
- Likely to be **more variable or more bursty**
  - **overflow processes**, in finite capacity systems, because they tend to occur in clumps, when the main system is overloaded.
  - **batch arrivals**, as in arrivals to amusement parks, arrivals at hospital ED because of accident or arrivals at ED that come by public transport (where customers may use resources as individuals).

Main Idea

Perform Statistical Tests with Arrival Data to see if the NHPP fits.

But how to do it?

Constructing Statistical Tests of an NHPP

1. Reduce to a statistical test of a PP by assuming that the arrival-rate function is piecewise-constant, and then focusing on the subintervals.
2. Reduce to statistical test of i.i.d. random variables with a specified distribution on each interval.
3. Use the Kolmogorov-Smirnov (KS) statistical test of i.i.d. random variables with a specified distribution.

Converting to a Kolmogorov Smirnov Test

1. Standard KS Test
   - Interarrival times of a PP are i.i.d. exponential random variables.
   - Use KS test for i.i.d. exponential variables on each interval
2. CU KS Test: Exploit Conditional Uniform (CU) Property to transform data and then use KS test (first key idea)
   - Given n arrivals A_k in [0,T]: A_k/T are n iid uniforms on [0,1]
   - No nuisance parameter: independent of arrival rate
   - Can combine data from different intervals and days
   - But note: **The scaled arrival times are i.i.d. uniforms, not the interarrival times!!**

The Kolmogorov-Smirnov (KS) Statistical Test (1)

Null Hypothesis, H0: We have a sample of size n from a sequence $\{X_k : k \geq 1\}$ of i.i.d. random variables with continuous CDF F

Alternative Hypothesis, H1: We have a sample of size n from a another (specified) sequence of random variables

The Kolmogorov-Smirnov (KS) Statistical Test (2)

The KS test is based on the absolute difference (a random variable, rv)

$$D_n \equiv \sup_x |F_n(x) - F(x)|,$$

where F is the postulated cdf in H_0, n is the sample size and $F_n(x)$ is the empirical CDF (another rv), i.e.,

$$F_n(x) \equiv \frac{1}{n} \sum_{k=1}^{n} 1_{X_k \leq x}, \quad -\infty < x < +\infty.$$

- Observe $D_n = \hat{D}_n$ for the data.
- Reject hypothesis H_0 if $\hat{D}_n > x_\alpha$, where $P(D_n > x_\alpha | H_0) = \alpha = 0.05$ (significance level)
- Compute p-value: $P(D_n > \hat{D}_n | H_0)$ (level for rejection)

Problem with the CU Test

KS test of NHPP using CU property has **very low power**

- Power: $P(\text{Reject}\, H0 | H1)$ (1 - type II error)
- Low power means that alternatives pass too easily!

Solve by applying KS test after data transformation (Apply KS test after producing new sequence of i.i.d. variables under H0)

- Log KS Test, from Brown (2005)
- Lewis KS Test, from Durbin (1961), Lewis (1965) and Kim and W (2014)

The Log KS Test from §3 of Brown et al. (2005)

Given n ordered arrival times $0 < T_1 < \cdots < T_n < t$ in $[0, t]$, let

$$X_j^{Log} \equiv -(n+1-j)\log_e\left(\frac{t - T_j}{t - T_{j-1}}\right), \quad 1 \leq j \leq n$$

- **Under H0:** If these rvs are obtained from a PP over $[0, t]$ using the CU property, then $\{X_j^{Log} : 1 \leq j \leq n\}$ are n i.i.d. mean-1 exponential rvs $\rightarrow$ KS Test
- The power is greater than the CU KS Test for most alternatives

The Lewis (1965) KS Test Based on Durbin (1961)

- $U_{(j)} = T_j/t$, $1 \leq j \leq n$ from $0 < T_1 < \cdots < T_n < t$ in $[0, t]$
ordered uniforms from the CU property
- $C_1 = U_{(1)}, C_j = U_{(j)} - U_{(j)}$ and $C_{n+1} = 1 - U_{(n)}$ and $C_{(j)}$ so that
$C_{(1)} < C_{(2)} < \cdots < C_{(n+1)}$ ordered intervals
- $Z_j = (n+2-j)(C_{(j)} - C_{(j-1)})$ scaled intervals

- **Under H0:** $(Z_1, \ldots, Z_n)$ is distributed the same as $(C_1, \ldots, C_n)$; $(S_1, \ldots, S_n)$ (where $S_j \equiv Z_1 + \cdots + Z_k$) is distributed the same as $(U_{(1)}, \ldots, U_{(n)}) \rightarrow$ KS Test
- The Lewis KS test has even more power!

Example: Different KS Tests Applied to an Alternative

Alternative: A renewal process with interarrival times having a hyperexponential CDF with $c_X^2 = Var(X)/(E[X])^2 = 2$

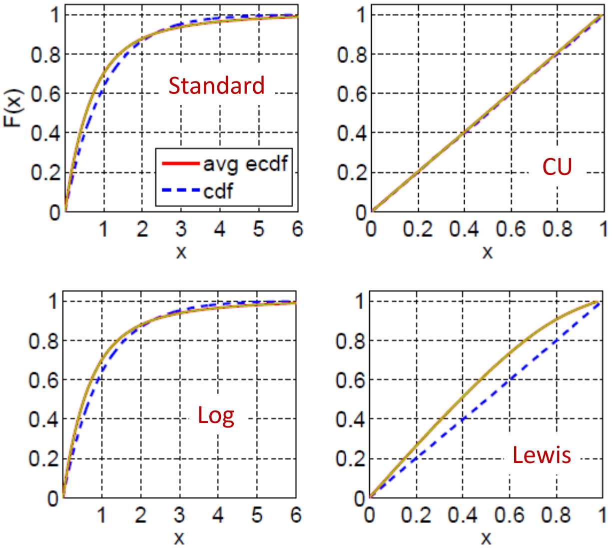
Simulation Experiment:

- Simulated 10^4 replications of 10^4 interarrival times
- Considered arrivals in $[10^3, 10^3 + 200]$ (stationary version)
- Apply KS tests to each replication w/ significance level $\alpha = 0.05$

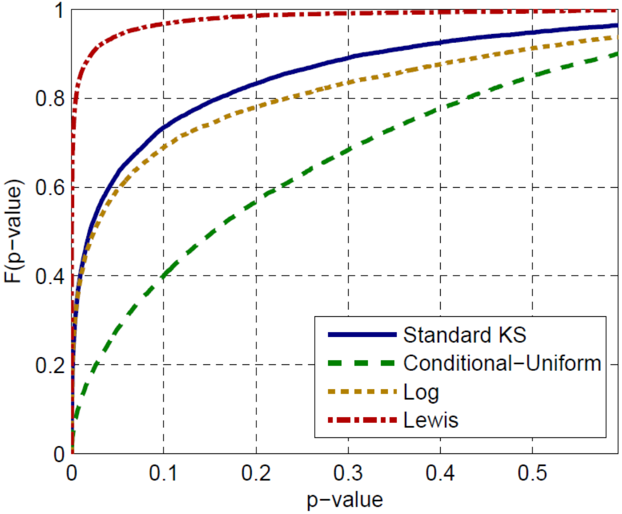
Table: Performance of alternative KS tests

KS test	Lewis	Standard	Log	CU
Power ($P(\text{Reject } H0 \mid H1)$)	0.94	0.63	0.51	0.28
Average p value	0.01	0.10	0.13	0.23

Insightful Plots 1: Average of ECDF over 10^4 Replications



Insightful Plots 2: ECDF of p-values over 10^4 Replications



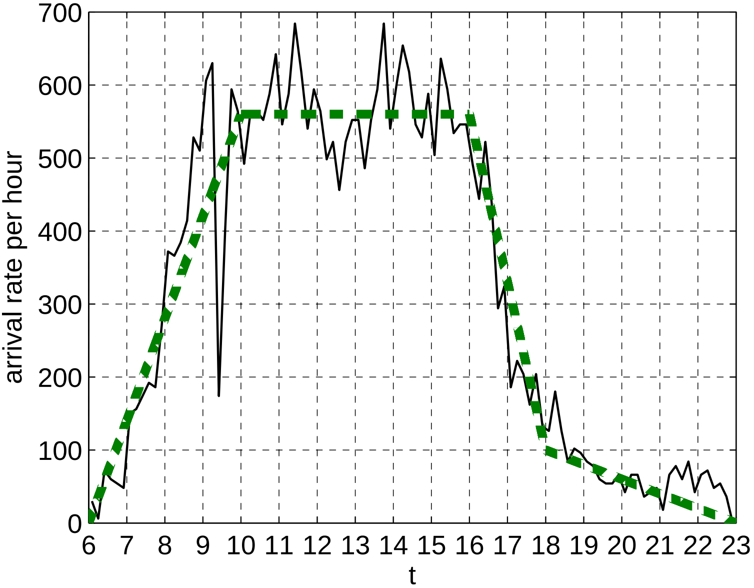
Applying the KS Tests to Service System Data

Three common features of arrival data we need to address (to avoid reaching the wrong conclusion):

1. data rounding, e.g., to seconds
2. choosing subintervals over which the rate varies too much
3. over-dispersion caused by combining data from multiple days that do not have the same arrival rate

In Kim and Whitt (2014, *M&SOM*) we use simulation to investigate how to address each problem, and apply our methods to call center and hospital arrival data.

Example: Call Arrivals at the U.S. Bank (slide 2)



Example: Lewis KS Test Applied to Call Center Data

Data: 30 days in April 2001 (more in KW2014, *M&SOM*)

Question: We observe that the arrival rate is nearly linear and increasing in the interval [7am, 10am]. We want to test whether the arrival process in [7am, 10am] is an NHPP.

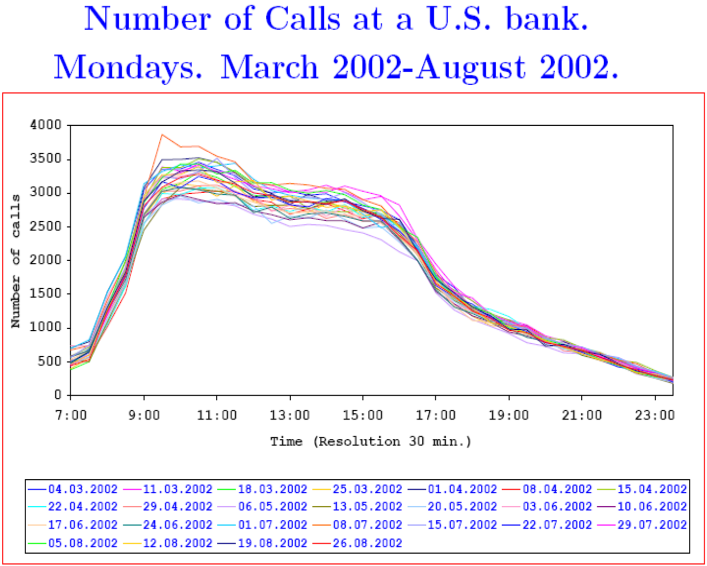
Table: Lewis KS test Results: The role of rounding and subintervals

subinterval length	Before unrounding		After Unrounding	
L (hours)	Avg p-value	# Pass	Avg p-value	# Pass
3	0.00 ± 0.01	1	0.04 ± 0.05	4
1.5	0.09 ± 0.08	7	0.26 ± 0.11	18
1	0.16 ± 0.08	15	0.48 ± 0.10	29
0.5	0.20 ± 0.09	18	0.51 ± 0.10	30

Overdispersion: Excessive Variability

- Overdispersion occurs in the arrival data if the arrival rate varies too much over successive days.
- It is not a problem if it can be predicted, e.g., by forecasting.
- It can be avoided in the KS tests by applying the CU transformation to different days separately and then combining the data afterwards.
- But overdispersion in arrival data is a genuine concern and needs to be examined.

Example from the U.S. Bank again



Estimate Day-To-Day Variation: 25 Mondays Overall

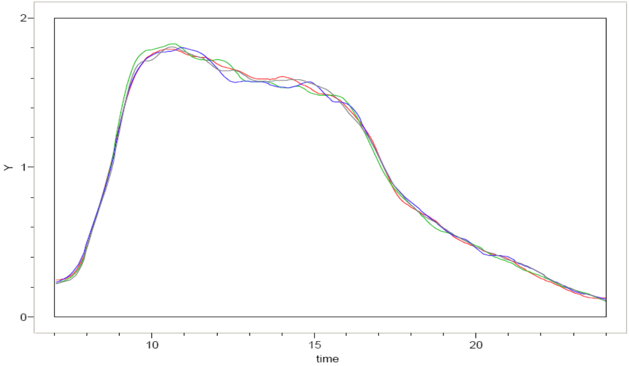
- Look at: **13:00-13:30**.
  - **Quick Rough Analysis**:
    - **range**: **[2500, 3200]**, **mean** ≈ 2850.
    - **5 STD DEV** ≈ 700, **STD DEV** ≈ 140.
    - **variance** $\approx 19,600 >> 2850$ **Too large!**.
    - **(The variance equals the mean in the Poisson distribution.)**
  - **Actual Data Analysis**:
    - **actual sample mean = 2,842**.
    - **actual sample variance = 24,539** $>>$ **2,842**. Thus, **Too large!**.

Separating Hourly Rate from the Daily Total: Normalize

Mondays

Plot shows (spline-smooth of)

$$\text{Normalized Arrivals}_{\text{given day}} = \frac{\text{Arrivals per Hour}_{\text{that day}}}{\text{Average}_{\text{that day}}\ (\text{arriv's. per hour})}$$



The variable **“time”** is on a 24 hour clock.

= 8/05/02
= 8/12/02
= 8/19/02
= 8/26/02

Summary (1 of 2)

Summary (1 of 2)

1. We have conducted extensive simulation experiments to study the power of alternative KS tests for a NHPP (arrival process model in service system models).

Summary (1 of 2)

1. We have conducted extensive simulation experiments to study the power of alternative KS tests for a NHPP (arrival process model in service system models).
2. The CU transformation allows us to apply the KS test to i.i.d. random variables and eliminates the nuisance parameter, but the CU test itself has very low power for non-exponential interarrival times.

Summary (1 of 2)

1. We have conducted extensive simulation experiments to study the power of alternative KS tests for a NHPP (arrival process model in service system models).
2. The CU transformation allows us to apply the KS test to i.i.d. random variables and eliminates the nuisance parameter, but the CU test itself has very low power for non-exponential interarrival times.
3. Our analysis strongly supports the approach proposed by Brown et al. (2005), but finds another related KS test proposed by Lewis (1965) consistently has even greater power.

Summary (1 of 2)

1. We have conducted extensive simulation experiments to study the power of alternative KS tests for a NHPP (arrival process model in service system models).
2. The CU transformation allows us to apply the KS test to i.i.d. random variables and eliminates the nuisance parameter, but the CU test itself has very low power for non-exponential interarrival times.
3. Our analysis strongly supports the approach proposed by Brown et al. (2005), but finds another related KS test proposed by Lewis (1965) consistently has even greater power.
4. We discussed three common features that need to be addressed when applying the KS tests to real arrival data.

Summary (2 of 2)

Three common features that need to be addressed when applying the KS tests to real arrival data:

1. data rounding

Summary (2 of 2)

Three common features that need to be addressed when applying the KS tests to real arrival data:

1. data rounding
2. choosing the subintervals to be treated as piecewise constant

Summary (2 of 2)

Three common features that need to be addressed when applying the KS tests to real arrival data:

1. data rounding
2. choosing the subintervals to be treated as piecewise constant
3. over-dispersion

   Over-dispersion appears to be the greatest practical concern in actual arrival processes. It can be tested for directly. If present at high levels, it requires new models and analysis techniques.

References

- Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. 2005. Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective. *Journal of the American Statistical Association* 100:3650.
- Durbin, J. 1961. Some Methods for Constructing Exact Tests. *Biometrika* 48(1):41-55.
- Kim, S.-H. and Whitt, W. 2014. Choosing Arrival Process Models for Service Systems: Tests of a Nonhomogeneous Poisson Process, *Naval Research Logistics* 61(1):66-90.
- Kim, S.-H. and Whitt, W. 2014. Are Call Center and Hospital Arrivals Well Modeled by Nonhomogeneous Poisson Processes?, *Manufacturing and Service Operations Management* 16(3):464-480.
- Kim, S.-H. and Whitt, W. 2015. The Power of Alternative Kolmogorov-Smirnov Tests Based on Transformations of the Data, *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 24(4):article 24.
- Lewis, P. A. W. 1965. Some Results on Tests for Poisson Processes. *Biometrika* 52 (1): 67-77.