Statistical Analysis of Service and Patience Data

From Service System Data to Performance Models

IEOR 4615, Service Engineering, Professor Whitt

Lecture 19, April 14, 2015

Toward a Stochastic Performance Model

- What do we **anticipate**?
  - We anticipate an **Erlang A** ($M/M/s+M$) **Model**.
- From last week: a **NHPP Arrival Process**.
- However, daily totals may not be Poisson.
  - over-dispersion in call centers
  - under-dispersion with appointment systems (outpatient clinics)
  - We may be able to reduce impact by **forecasting**:
  - We **estimate the arrival rate function**.
- For staffing, we assume a **piecewise-constant arrival rate function**.
- Focus on individual hours or half hours: **stationary model**.

Toward a Stochastic Performance Model

- What do we **anticipate**?
  - We anticipate an **Erlang A** **(**$M/M/s+M$**)** **Model**.
- Today consider **service-time and patience-time distributions**.
  - Assuming that these are i.i.d. exponential variables,
  - we **estimate mean service time** $1/\mu$
  - and we **estimate abandonment rate** θ. **(censored data)**.
- Assuming more general distributions: more complicated

Estimating the Service and Patience Parameters

- Estimate **mean service time** $1/\mu$ of each customer.
  - Observe n service times $S_1, S_2, \ldots$; let $1/\hat{\mu} = n^{-1} \sum_{k=1}^{n} S_k$.
- Estimate **abandonment rate** θ for each customer from queue.
  - Problem: **censored data**. (We do not observe most patience times.)
  - Use $P(Ab|Wait > 0) = \theta E[Wait|Wait > 0]$ (explained on next slide).
    - Observe outcomes for n arrivals that must wait (join the queue).
    - Observe positive waiting times $W_1, W_2, W_3, \ldots$.

      Estimate $E[Wait|Wait > 0]$ by $\bar{W}_n \equiv n^{-1} \sum_{k=1}^{n} W_k$.
    - Observe number $N_{ab}(n)$ of these that abandon.

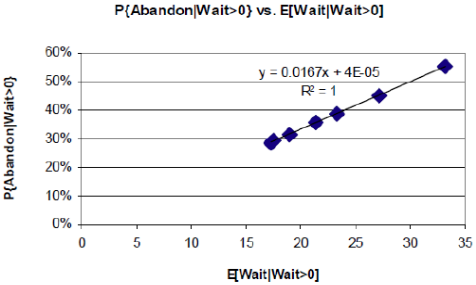
      Estimate $P(Ab|Wait > 0)$ by $\bar{N}_{ab} \equiv N_{ab}(n)/n$.
    - Estimate θ by $\hat{\theta} = \bar{N}_{ab}/\bar{W}_n$,

      using $\theta = P(Ab|Wait > 0)/E[Wait|Wait > 0]$ from above.

Justifying the Abandonment Rate Formula

- Why $P(Ab|Wait > 0) = \theta E[Wait|Wait > 0]$ in the Erlang-A model?
- **Use structure of the Erlang A model.**
- **Apply Little's Law.**
  - Let Q and W be steady-state number in queue and waiting time.
  - $P(Ab) = \frac{\text{abandonment rate}}{\text{arrival rate}} = \frac{\sum_{k=1}^{\infty} P(Q=k)k\theta}{\lambda} = \frac{E[Q]\theta}{\lambda} = \theta E[W]$.
  - Extend to conditioning upon positive wait:
    1. $P(Ab) = P(Ab|W > 0)P(W > 0)$ and $E[W] = E[W|W > 0]P(W > 0)$,
    2. So $P(Ab|W > 0)P(W > 0) = P(Ab) = \theta E[W] = \theta E[W|W > 0]P(W > 0)$.
    3. Then divide by $P(W > 0)$ on both sides.

Apply Regression to Multiple Samples



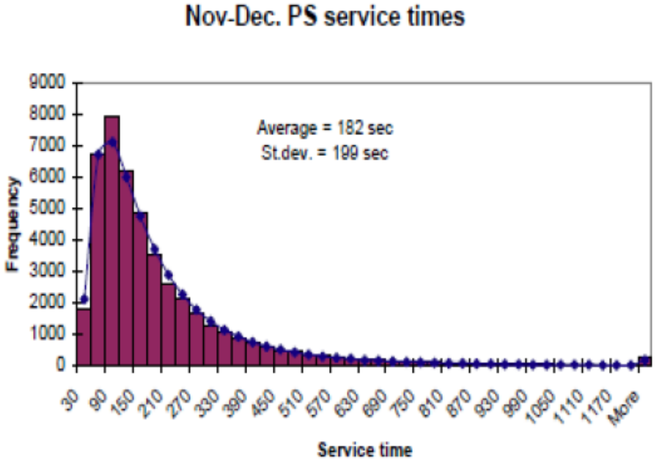
and we get

$$P\{Abandon \mid Wait > 0\} = 4.35 \cdot 10^{-5} + 0.01666 E[Wait \mid Wait > 0]$$
$$\approx 0.01666 E[Wait \mid Wait > 0]$$

$$\Rightarrow \quad Average\ Patience = \frac{1}{0.01666} = 60.00817 \text{ sec}$$

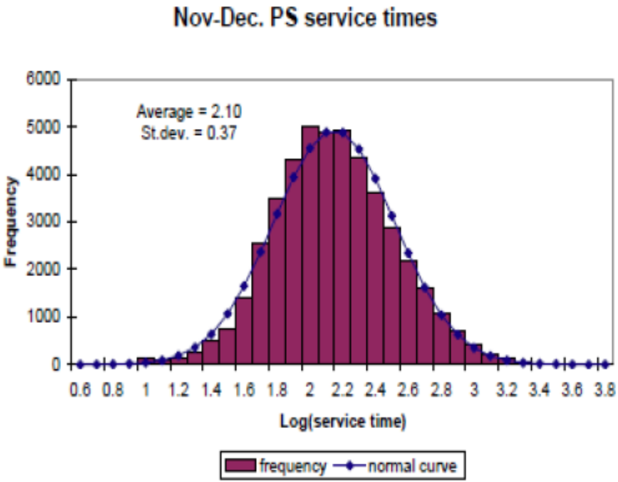
What Do We See in the Data?

Histogram of Banking Call Center Service Times



Is Logarithm Normally Distributed?

Histogram of Logarithm of Banking Call Center Service Times



Lognormal Distribution

- X has a lognormal distribution if $\log_e X$ has a normal distribution:
  - $P(\log_e X \leq y) = P(N(\mu, \sigma^2) \leq y)$ or X is distributed as $e^{N(\mu,\sigma^2)}$.
  - cdf: $P(X \leq x) = P(\log_e X \leq \log_e x) = P(N(\mu, \sigma^2) \leq \log_e x)$.
- Other characteristics:
  - pdf: $f_X(x) = \frac{d}{dx} P(X \leq x) = \frac{d}{dx} P(N(\mu, \sigma^2) \leq \log_e x)$
    $= \frac{f_{N(\mu,\sigma^2)}(\log_e x)}{x} = \frac{1}{\sqrt{2\pi\sigma^2}x} e^{-(\log_e x - \mu)^2/2\sigma^2}$.
  - moments: $E[X^k] = E[e^{kN(\mu,\sigma^2)}]$, but mgf: $E[e^{tN(\mu,\sigma^2)}] = e^{\mu t + \sigma^2 t^2/2}$ (p. 64 of Ross 2010) so that first two moments are $E[X] = e^{\mu + \sigma^2/2}$ and $E[X^2] = e^{2\mu + 2\sigma^2}$, so that $c_X^2 = \frac{E[X^2]}{E[X]^2} - 1 = e^{\sigma^2} - 1$.
  - median $= e^{\mu}$, mode $= e^{\mu - \sigma^2}$, so mean $>$ median $>$ mode, somewhat heavy tail.

Median of Lognormal Distribution

- $P(\log_e X \le y) = P(N(\mu, \sigma^2) \le y)$.
- median $= e^{\mu}$, Proof:

$$\begin{aligned} P(X \le m) &= 1/2 \quad \text{for} \quad m = \quad \text{median}, \\ P(\log_e X \le \log_e m) &= 1/2 \\ P(N(\mu, \sigma^2) \le \log_e m) &= 1/2 \\ \text{so that} \quad \log_e m &= \mu \quad \text{and thus} \quad m = e^{\mu}. \end{aligned}$$

Mode of the Lognormal Distribution

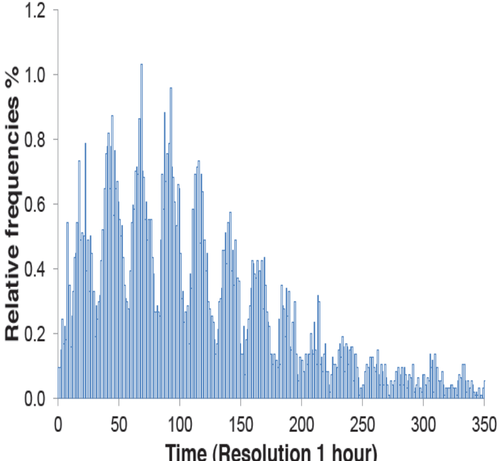
- cdf: $P(X \leq x) = P(\log_e X \leq \log_e x) = P(N(\mu, \sigma^2) \leq \log_e x)$.
- pdf: $f_X(x) = \frac{d}{dx}P(X \leq x) = \frac{d}{dx}P(N(\mu, \sigma^2) \leq \log_e x)$
  $= \frac{f_{N(\mu,\sigma^2)}(\log_e x)}{x} = \frac{1}{\sqrt{2\pi\sigma^2}x}e^{-(\log_e x-\mu)^2/2\sigma^2}$.
- mode $= e^{\mu-\sigma^2}$, Proof:

$$\frac{d}{dx}f_X(x) = 0 \quad \text{for} \quad f_X(x) = \frac{Ae^{-u(x)}}{x} \quad \text{and} \quad u(x) = \frac{(\log_e(x) - \mu)^2}{2\sigma^2}$$

$$\begin{aligned} 0 &= \frac{-Ae^{-u(x)}}{x^2} + \frac{-A\dot{u}(x)e^{-u(x)}}{x} \\ &= 1 + \dot{u}(x) = 1 + \frac{\log_e(x) - \mu}{\sigma^2} \\ &\quad \text{or} \quad \log_e(x) - \mu = -\sigma^2 \end{aligned}$$

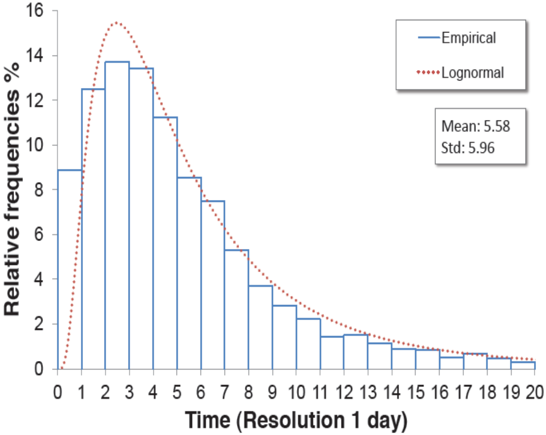
Strange Service Times

Length-of-Stay (LOS) at the Internal Wards of a Hospital: Mixture (of Skewed-Normals), in *Hours*



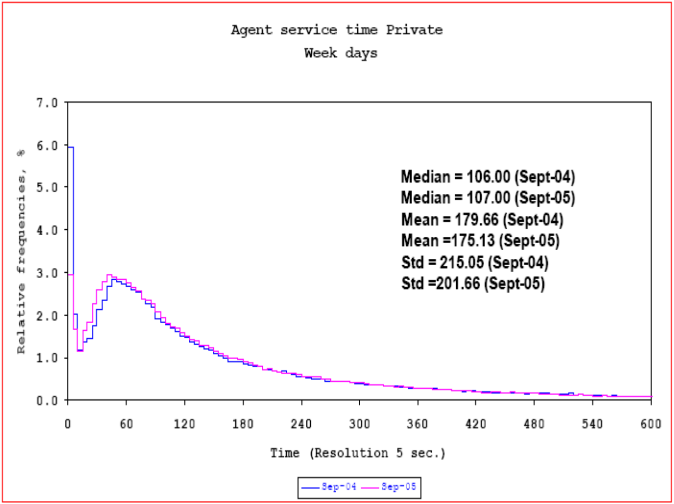
Measure in Days

Length-of-Stay (LOS) at the Internal Wards of a Hospital: LogNormality, in *Days*



Strange Service Times: Israeli Telecom

IL Telecom: Dynamics of the distribution of agent service time for Private calls



Diagnosis: Look at Individual Agents

Table 52: Number of calls handled by an agent

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
AVI	0	0	0	1117	2208	2019	2789	2710	1417	2026	2523	2395
AVNI	1493	1736	642	539	1786	2219	2092	2392	1156	1888	1988	2136
BASCH	999	1164	1708	1155	982	906	858	2185	1973	1055	1326	1242
BENSION	1283	1135	0	1053	1108	1016	1682	1298	1076	1303	1546	1176
DARMON	309	515	633	519	577	436	309	370	297	194	425	128
DORIT	696	1047	0	811	546	862	750	2228	1319	1384	1640	1605
ELI	387	508	777	447	560	436	395	458	416	363	502	352
GELBER	333	143	510	427	859	281	386	332	67	179	165	269
GILI	668	614	1155	803	1108	974	418	0	355	456	412	298
KAZAV	1995	1693	1240	1451	1731	2251	1737	1168	729	1570	1047	2038
MEIR	0	0	0	0	0	0	127	344	318	280	406	454
MORIAH	1360	1223	1591	1351	1866	1980	2416	2152	1526	1940	1793	515
PINHAS	79	40	359	244	31	311	422	241	143	105	51	63
ROTH	0	0	397	1292	1928	1967	1831	1749	1625	1914	1458	1038
SHARON	1985	1674	2780	1938	2563	2657	2537	2875	1803	1935	2532	2140
STEREN	0	1043	2294	1516	2163	2231	1423	2455	1672	709	2375	2568
TOVA	1923	1679	1562	1059	1464	1389	1890	1811	1361	1971	941	0
VICKY	895	0	0	0	1006	1378	1415	1674	1472	1582	1641	1990
YIFAT	1312	1901	1745	1305	1464	1076	780	90	1137	1315	0	0
YITZ	1771	1791	1402	1203	1355	1367	1009	69	705	1743	2420	2353
ZOHARI	891	1144	1398	1148	1479	1450	980	1494	1423	1359	1504	1094
Z2ARIE	0	0	0	0	0	0	0	56	225	315	432	534
Z2ELINOR	0	0	0	0	0	0	0	45	352	288	222	310
Z2EYAL	0	0	0	0	0	0	0	95	331	428	579	618
Z2IFAT	0	0	0	0	0	0	0	94	260	314	215	0
Z2LIOR	0	0	0	0	0	0	0	84	250	136	126	138
Z2NIRIT	0	0	0	0	0	0	0	116	327	474	387	545
Z2OFERZ	0	0	0	0	0	0	0	71	311	260	242	334
Z2SPIEGEL	0	0	0	0	0	0	0	71	311	260	153	322

Diagnosis: Look at Individual Agents

Table 53: Number of calls with short service time

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
MORIAH	233	230	356	290	614	695	865	597	490	455	4	1
AVI	0	0	0	47	111	144	295	221	121	76	35	26
AVNI	11	13	4	5	6	25	16	18	4	8	8	11
DARMON	2	11	8	9	10	7	1	0	1	1	0	0
ELI	9	7	10	12	22	18	15	4	8	3	6	5
KAZAV	57	40	48	44	48	63	40	27	15	18	4	6
MEIR	0	0	0	0	0	0	1	8	3	1	2	1
PINHAS	3	0	58	25	4	14	11	6	8	1	0	0
ROTH	0	0	10	10	36	21	43	25	32	31	3	6
SHARON	58	49	86	52	67	78	66	63	38	23	43	49
TOVA	52	163	269	132	231	193	100	109	207	190	6	0
ZOHARI	4	8	12	22	17	20	9	14	5	7	10	7

Kolmogorov-Smirnov (KS) Tests for Service Times

- **H0: $\{S_k\}$ iid sequence with $P(S_k \leq x) = F(x)$** with F continuous.
- Use ECDF $\hat{F}_n(x) \equiv n^{-1} \sum_{k=1}^{n} 1_{\{S_k \leq x\}}$ and $\hat{D}_n \equiv \sup_x |\hat{F}_n(x) - F(x)|$.
  - Observe n service times S_k. Under H0, $Y_k = F(S_k)$ are iid **uniforms** on $[0, 1]$. ($P(F(S_k) \leq x) = P(S_k \leq F^{-1}(x)) = F(F^{-1}(x))$.)
  - Under H0, $Z_k = -\log(1 - Y_K)$ iid **exponentials**. (same proof)
- **Alternative KS Tests** with **H0: $\{Y_k\}$ iid uniform on $[0, 1]$, $F(x) = x$.**
  - **standard KS:** Reject if $P(D_n > \hat{D}_n | H0) > \alpha = 0.05$.
  - **Durbin:** Sort to get ordered uniforms, do as before (last class).
  - **Exp+CU:** With exponentials Z_k and $S_k = Z_1 + \cdots Z_k$, apply Conditional Uniform (CU) to get S_k/S_n ordered uniforms on $[0, 1]$.
  - **Exp+CU+Durbin (Kim&W 2015):** Do Durbin after Exp+CU, as in Lewis.

Example: Different KS Tests Applied to an Alternative

Simulation Experiment: Apply the KS test of iid mean-1 exponential service times to the **alternative:** iid mean-1 H_2 (hyperexponential, mixture of two exponentials) service times with $c_X^2 = Var(X)/(E[X])^2 = 2.0$.

Table: **Performance of alternative KS tests** of i.i.d. mean-1 exponential variables for the sample size $n = 200$ with significance level $\alpha = 0.05$: the case of i.i.d. H_2 interarrival times having $EX = 1$ and $c_X^2 = 2$, based on 10^4 replications.

KS test	Exp+CU+Durbin	Standard	Exp+CU	Durbin
KS test	Kim&W	Standard		
Power	0.93	0.64	0.28	0.14
Average p value	0.02	0.09	0.24	0.40

Previous Example: KS Tests of a Poisson Process

Simulation Experiment: Apply the KS test to the **alternative:** a non-Poisson renewal process with interarrival times having an H_2 (hyperexponential) CDF (mixture of two exponentials) with a squared coefficient of variation $c_X^2 = Var(X)/(E[X])^2 = 2.0$.

Table: **Performance of alternative KS tests** of a rate-1 Poisson process for the time interval $[0, 200]$ with significance level $\alpha = 0.05$: the case of a renewal process with H_2 interarrival times having $c_X^2 = 2$, based on 10^4 replications..

KS test	Lewis	Standard	Log	CU
Power	0.94	0.63	0.51	0.28
Average p value	0.01	0.10	0.13	0.23

References

- L. Brown et al. Statistical Analysis of a Telephone Call Center: A Queueing Science Perspective. *Journal of the American Statistical Association* (JASA) 100 (2005) 36-50.
- J. Durbin. Some Methods for Constructing Exact Tests. *Biometrika* 48 (1961) 41-55.
- S-H. Kim and WW. The Power of Alternative Kolmogorov-Smirnov Tests Based on Transformations of the Data. *ACM Transactions on Modeling and Computer Simulation* (TOMACS), 2015, forthcoming.
- P. A. W. Lewis. Some Results on Tests for Poisson Processes. *Biometrika* 52 (1965) 67-77.

Exponential Distribution:

Density: $f(x) = \lambda e^{-\lambda x}, \ x \geq 0,$

Mean: $\mathrm{E}[X] = \lambda^{-1},$

Variance: $Var(X) = \lambda^{-2},$

Coefficient of Variance: $C_v = \dfrac{SDV(X)}{\mathrm{E}[X]} = 1,$

Median: $\lambda^{-1} ln 2.$

Check that the Parallel Server Model Makes Sense

One Queue with Multiple Servers Working in Parallel