Irregular Simulation: Input Modeling and Applications

W. David Kelton

Professor, Dept. of Operations, Business Analytics, and Information Systems, University of Cincinnati

Visiting Professor, Dept. of Operations Research, Naval Postgraduate School

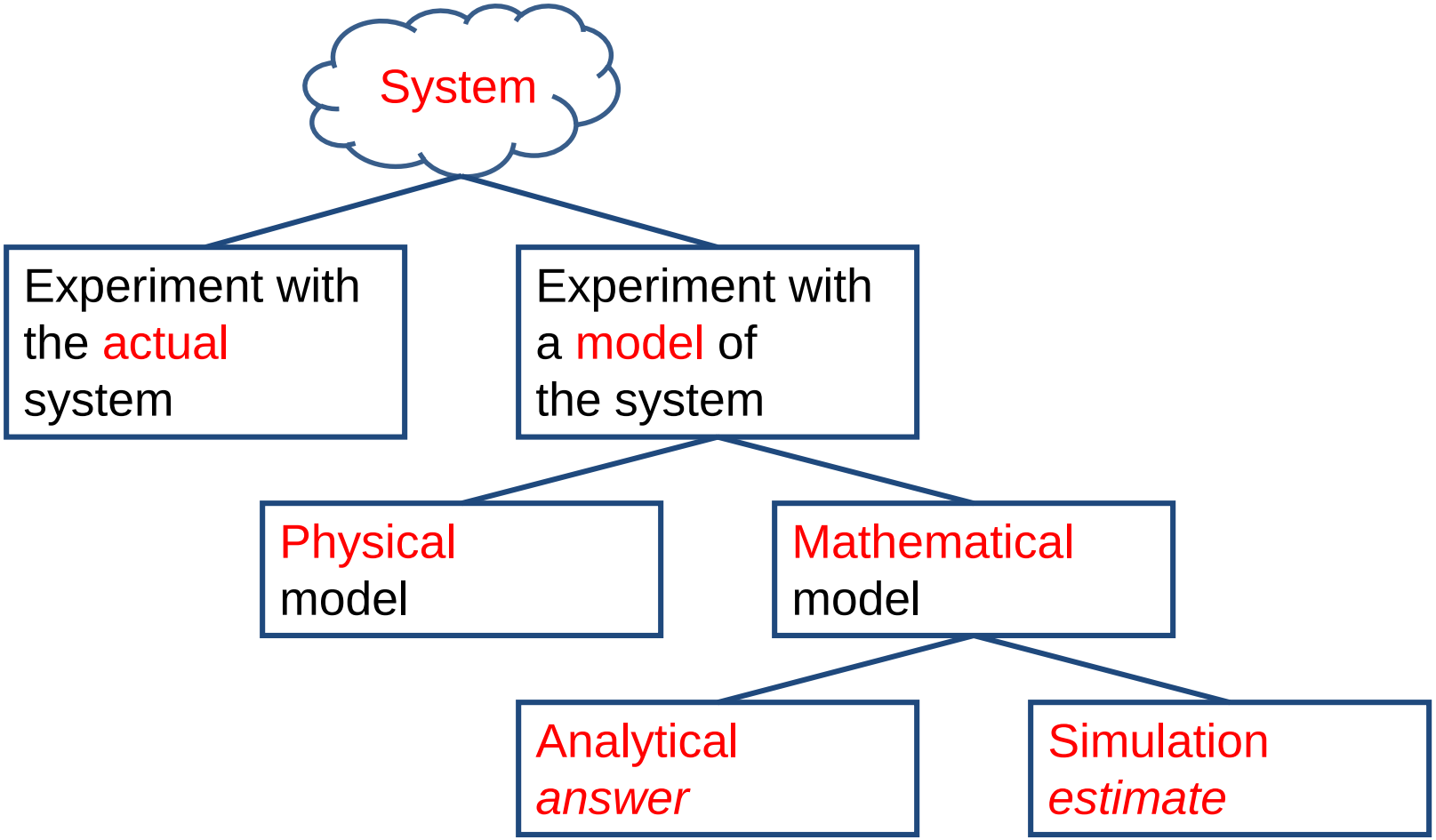
david.kelton@uc.edu

http://www.cba.uc.edu/faculty/keltonwd

Outline

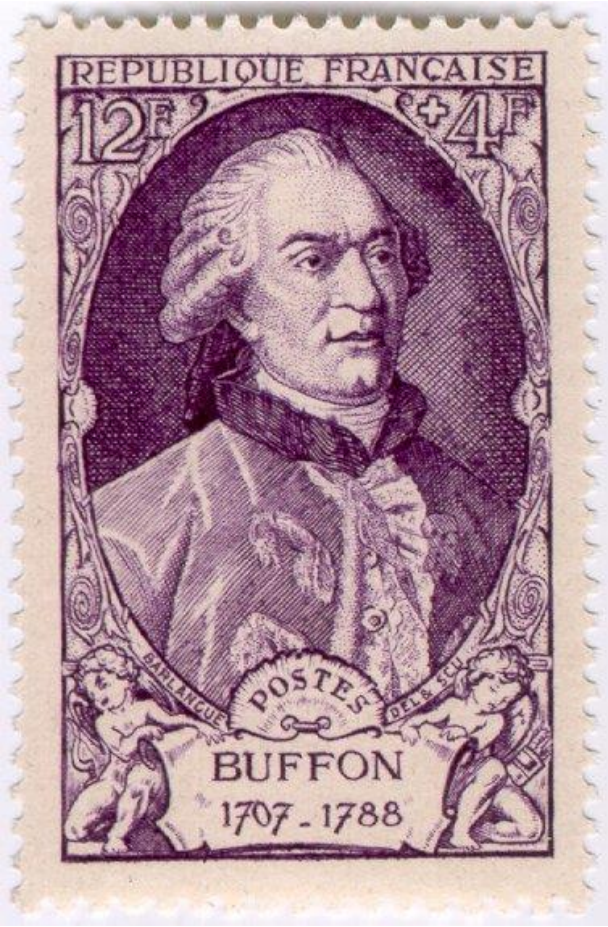
- Systems and Studying Them
- Why Not, Why, How Not, and How to Simulate
- Probability Inputs Matter
- Lanchester Combat Differential Equations
- US West-Coast Container-Port Operations
- We Need Better Simulation Software!
- Conclusions

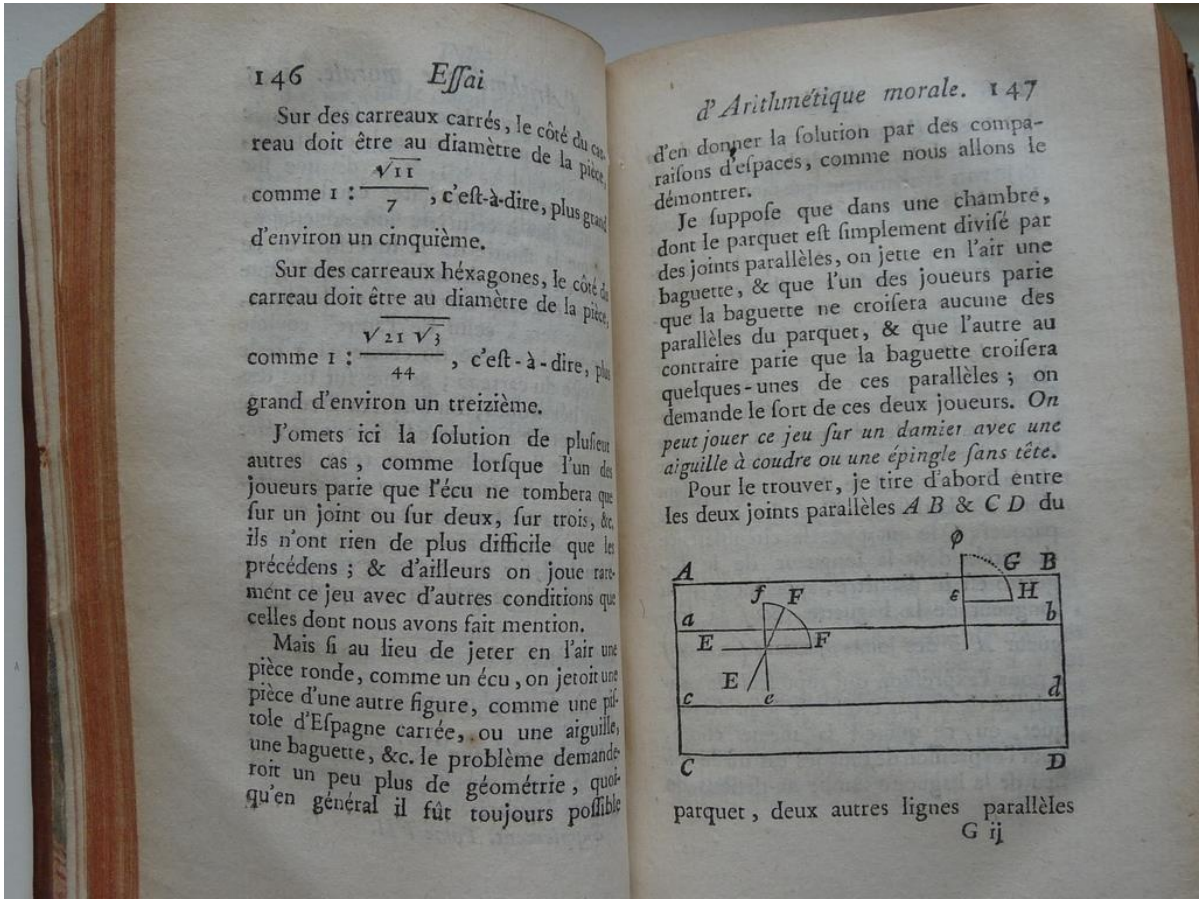
Systems and Studying Them



"When all else fails …Method of last resort"
– Harvey Wagner, 1969

Why Not, Why, How Not, and How to Simulate





146 *Essai*

Sur des carreaux carrés, le côté du carreau doit être au diamètre de la pièce, comme 1 : $\frac{\sqrt{11}}{7}$, c'est-à-dire, plus grand d'environ un cinquième.

Sur des carreaux héxagones, le côté du carreau doit être au diamètre de la pièce, comme 1 : $\frac{\sqrt{21\sqrt{3}}}{44}$, c'est-à-dire, plus grand d'environ un treizième.

J'omets ici la solution de plusieurs autres cas, comme lorsque l'un des joueurs parie que l'écu ne tombera que sur un joint ou sur deux, sur trois, &c. ils n'ont rien de plus difficile que les précédens ; & d'ailleurs on joue rarement ce jeu avec d'autres conditions que celles dont nous avons fait mention.

Mais si au lieu de jeter en l'air une pièce ronde, comme un écu, on jetoit une pièce d'une autre figure, comme une pistole d'Espagne carrée, ou une aiguille, une baguette, &c. le problème demanderoit un peu plus de géométrie, quoiqu'en général il fût toujours possible

d'Arithmétique morale. 147

d'en donner la solution par des comparaisons d'espaces, comme nous allons le démontrer.

Je suppose que dans une chambre, dont le parquet est simplement divisé par des joints parallèles, on jette en l'air une baguette, & que l'un des joueurs parie que la baguette ne croisera aucune des parallèles du parquet, & que l'autre au contraire parie que la baguette croisera quelques-unes de ces parallèles ; on demande le sort de ces deux joueurs. *On peut jouer ce jeu sur un damier avec une aiguille à coudre ou une épingle sans tête.*

Pour le trouver, je tire d'abord entre les deux joints parallèles *A B* & *C D* du

parquet, deux autres lignes parallèles

G ij

Why Not To Simulate

1. Have a valid mathematical model of the system
2. This valid model is simple enough to admit an analytical answer to your questions
3. Really, Really, ***REALLY*** believe 1 and 2 are true (not just convenient)

Example: M/G/1 queue, runs forever

E(waiting time in queue) =

$$\lambda \frac{\mathrm{Var}(S) + [\mathrm{E}(S)]^2}{2[1 - \lambda \mathrm{E}(S)]}$$

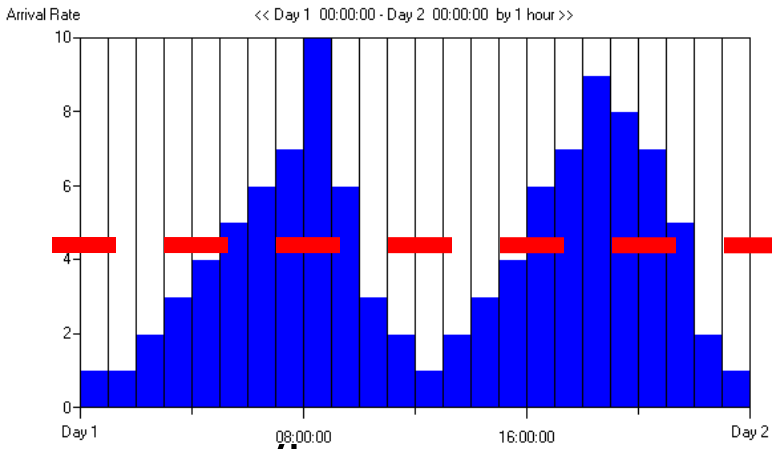
λ = arrival rate (exponential)

S = service-time random variable

Why To Simulate

- If any of 1, 2, or 3 on the preceding slide is false
- Main issue: model *validity*
  - Many systems are complicated, so *valid* models will also likely be complicated, analytically intractable
  - Could make simplifying *assumptions* on model to make it fit into an analytical-solution mold
    - Model validity?
    - Value of an exact answer?
    - Type III error; looking where the light is
  - Hard to quantify "how invalid" a model is
    - (Relatively) easy to quantify imprecision of simulation results
    - If I don't like the answer I can do something about it ...

Why To Simulate – Example



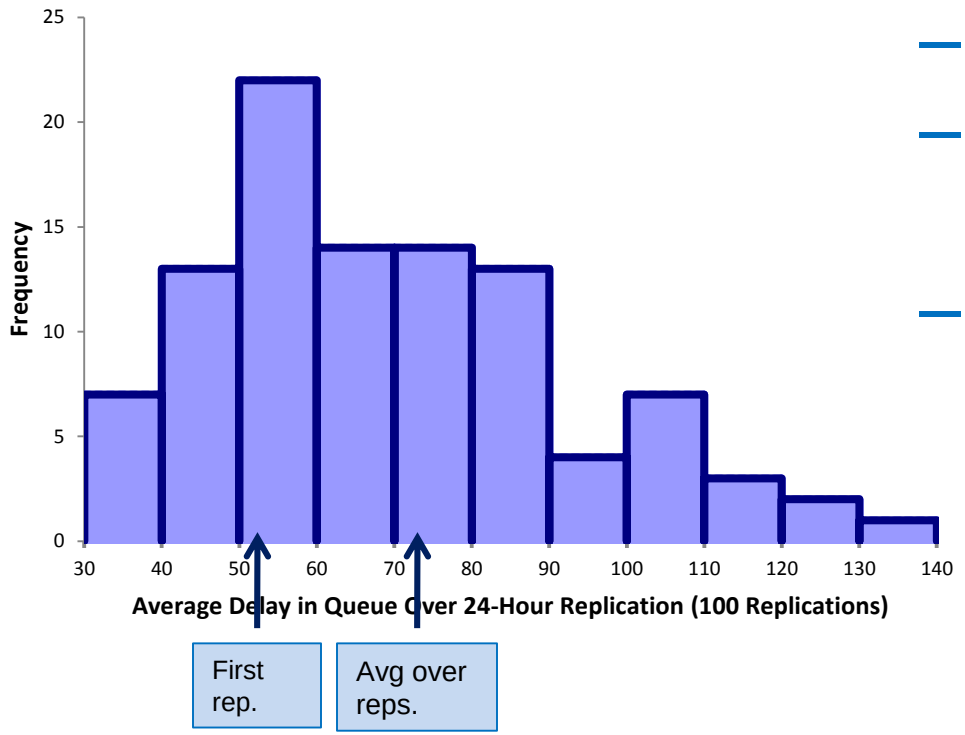
- Incoming message traffic, nonstationary Poisson process, rate function (messages/hour)
  - *Average* arrival rate over day = 4.375 messages/hour, but that shouldn't be used ...
- Service times Unif [6, 18] minutes, single server
- Start empty/idle, operate for 24 hours
- Outputs: average message delay in queue, (time-)average number of messages in queue
- Message processing.doe

Why To Simulate – Example Results

Approach	Model	Modeling errors	E(Delay in queue)	E(Num. in queue)
Simulate 100 reps.	NS/U/1	None	70.1 ± 4.6	5.5 ± 0.4
	M/U/1	Flat arrival rate	30.6 ± 3.4	2.4 ± 0.3
Analytical	M/U/1	Flat arrival rate, duration	45.5 ± 0	3.3 ± 0
	M/M/1	Flat arrival rate, duration, service-time distrib.	84.0 ± 0	6.1 ± 0

How Not To Simulate

- Make a run, get The Answer
  - NSPP/U/1, rep. 1, avg. delay in queue = 50.79 mins.
- Histogram of 100 such numbers, 100 reps.:

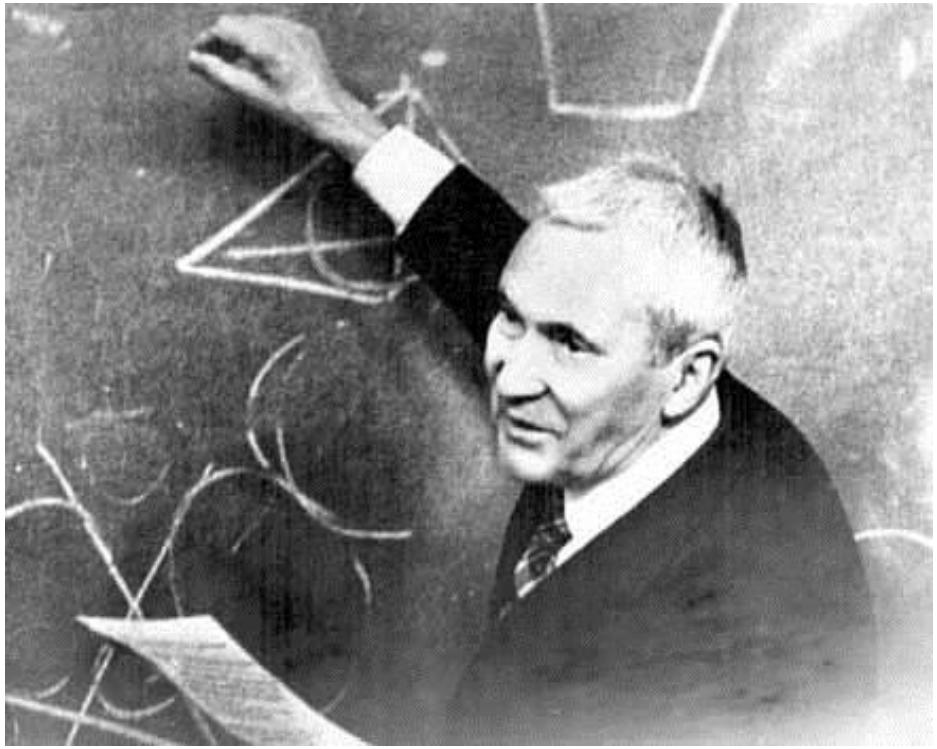


  - Simulation is an *experiment*
  - Variation is present, must measure, reduce if necessary
  - Issues: number of reps., rep. length (maybe), variance estimation

How To Simulate

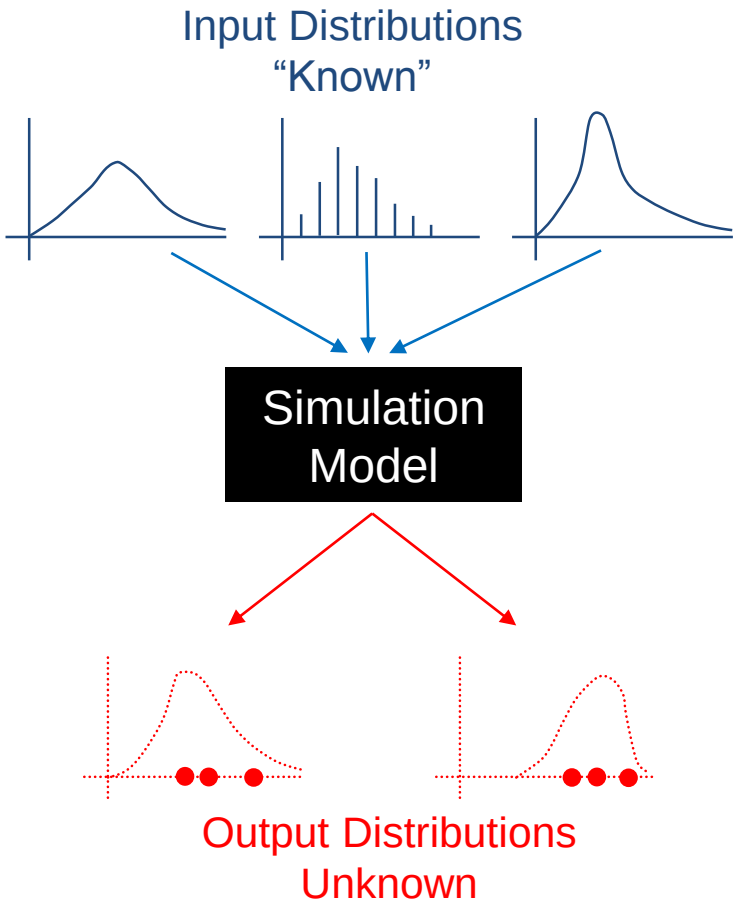
- Make sure you need to simulate
- Build a valid, verified model
  - Right level of detail ... more detail is not always better
  - Programming language (perhaps augmented) vs. simulation software vs. template
- Time frame of simulation: long-run or terminating
- Variance estimation – beware Stat 101 methods
  - Number of replications for terminating
  - Number of replications or run length for long-run
- Statistical analysis of simulation output
- Design of simulation experiments

Probability Inputs Matter



Main Disadvantage of Simulation

- Don't get exact answers, only approximations, *estimates*
- Stochastic simulations produce random output
  - Statistical design, analysis of simulation experiments
  - Exploit: noise control, replicability, sequential sampling, variance-reduction techniques
  - Catch: "standard" statistical methods seldom work … *simulation output analysis*



RIRO

Focus on Input-Modeling Side

- Options to *represent* and *generate* uncertain inputs to simulation models
- Default ways of doing this in simulation software may not always be appropriate
  - Can lead to inaccuracy, imprecision/inefficiency
  - But if done right, can improve model validity, precision/efficiency
  - The trick is to be aware of the issues and possibilities

The Declaration of Independence (a.k.a. The Titanic Assumption)

- Most random inputs in dynamic simulations:
  - *Independent* and *Identically Distributed* (IID)
    - Separate draws from *this* distribution are independent of each other as the simulation progresses
    - *This* distribution does not change in any way as the simulation progresses
    - Draws from *this* distribution are independent of draws from all *other* distributions in the model
    - TandemProcess-Indep.doe
      - Three sources of randomness (interarrivals, two service times)
    - It's hard to do otherwise in most simulation software!

The Declaration of Independence (cont'd.) Across Inputs

- But it's easy to imagine models where correlated inputs would be needed
  - Same model, but Cor(ServiceTime$_1$, ServiceTime$_2$) = 0.9
    - TandemProcess-Correl.doe
    - Had to "program" generation of correlated service times in Assign module … trivariate reduction for correlated gammas
  - ***It matters***: 1000 five-day replications, 95% CIs

	Independent	Correlated
Avg. total time in system (minutes)	88.0 ± 2.0	79.0 ± 1.9
Avg. total WIP	8.9 ± 0.2	8.0 ± 0.2

The Declaration of Independence (cont'd.) Within an Input

- Possibility of *auto*correlation within an input process
  - Telecommunications system, packets are arriving
  - Large (small) packets tend to be followed by other large (small) packets … positive autocorrelation of packet sizes
  - Ignoring autocorrelation, and modeling packet sizes as independent, could lead to understatement of system congestion
    - Many large packets in a row $\Rightarrow$ great traffic buildup
    - Queues are slower to clear out than they are to build up

Cautions on the Normal Distribution

- Normal distribution is popular, comfortable
  - But in dynamic simulations, many input quantities must be positive or at least non-negative
    - Service times; Interarrival times; Up/down times
  - Normal distribution always has infinite left tail, so there's always a chance of generating negative values, maybe causing your model to go nuts, have big effect
  - NegativeNormals.xls: negative values are "lit up"
    - Left block of 1000: μ is 3 standard deviations > 0
    - Right block of 1000: μ is 4 standard deviations > 0

Using Observed Data

- Usually (hopefully), have field data on inputs
  - Interarrivals, service times, up/down times, …
- Ways to use observed data to specify inputs … each has pros, cons:
  1. Feed observed data directly into simulation to "drive" it directly … *trace-driven*
  2. Fit "standard" distributions to observed data
  3. Compromise: use an *empirical* distribution
- Radical proposals:
  - Mixed empirical-exponential distribution
  - Link simulation model directly back to data set

Assigning Random Numbers

- All stochastic simulations rely on an underlying random-number generator
  - Algorithm to generate what appear to be IID random variates distributed continuously uniformly on [0, 1]
- Without a good RNG, stochastic simulation is impossible … what is "good?"
  - Long cycle length (period)
  - Good statistical properties – uniformity, independence
  - Segmented into (long) separate, contiguous subsegments called *streams*
  - Unfortunately, some RNGs still in use fall short

Assigning Random Numbers (cont'd.)

- So assuming a good RNG, modeler should think about something other than the simulation software's default method of FCFS single-stream RN assignment
  - *Synchronize* RN use across model:
    - When comparing two or more competing scenarios, want to subject them to the same external random "shocks" so observed differences can be attributed to scenario differences, not to random bounce
    - Reduces variability in estimates of differences between scenarios

Assigning Random Numbers (cont'd.)

- Approaches to RN synchronization
  - *Faucets*: Use a separate RN stream for each source of randomness in model
    - Might not guarantee full, perfect synchronization if the number of RNs from a source differs across scenarios
    - Could be (but isn't) automated in simulation software
  - *Body art*: Pre-assign (to entity attributes) all random variates that *might* be used downstream to each arriving entity
    - Could entail large memory requirements
    - Might accomplish better synchronization than faucets

Lanchester Combat Differential Equations

Lanchester's Combat Models

- $x(t)$, $y(t)$ are size of blue, red forces at time t
- Modern: force depletion depends directly on opponent's force size and effectiveness:

$$\frac{dx}{dt} = -ay(t) \qquad \frac{dy}{dt} = -bx(t)$$

- Ancient: force depletion depends directly on opponent's force size and effectiveness, and on your force size:

$$\frac{dx}{dt} = -ax(t)y(t) \qquad \frac{dy}{dt} = -bx(t)y(t)$$

Lanchester in Arena

- Arena has Runge-Kutta numerical differential-equation *numerical* solving ability
- Specify differential equations via VBA
- Elements panel to set parameters, conditions
- Generalization of Lanchester (must simulate)
  - Both sides receive re-enforcements
  - Timing, size of re-enforcements is random
  - Trace of $x(t)$, $y(t)$ solutions is random (as is who wins)

US West-Coast Container-Port Operations

US West-Coast Container-Port Ops

- MS Thesis, NPS OR student
- Principal west-coast container-port operations
  - Huge economic impact, national-security implicatons
  - 96-hour notice of arrival, intended port
  - Unload containers at port, landside yard storage, load a "pile" of containers onto truck/rail for transport to one-digit ZIP codes
    - Divide US into 10 regions, population center in each
  - Costs for demurrage, operations
  - Ignore outbound freight, empties

Transportation Security Incident (TSI)

" …a security incident that results in a significant loss of life, environmental damage, transportation system disruption, or economic disruption in a particular area."

Maritime Transportation Security Act, 2002

"In the aftermath of a Transportation Security Incident, the recovery of critical infrastructures, resumption of the Marine Transportation System, and restoration of communities within the affected area must all occur simultaneously and expeditiously."

The White House, 2005

Modeling a TSI

- Take one or more ports out, completely or partly
  - Attack, strike/lockout, earthquake, ...
    - Ten-day lockout in fall 2002, use as a benchmark
  - Estimate downtime, how port comes back up
  - During downtime, inbound vessels must divert to another port
    - Choose based on draft of vessel, look at congestion in other ports, pick the one that will likely result in soonest unload
    - Mimics shippers' own economic self-interest
  - Measure economic, operational impact
  - Identify ports that are particularly critical to operations and to the economy

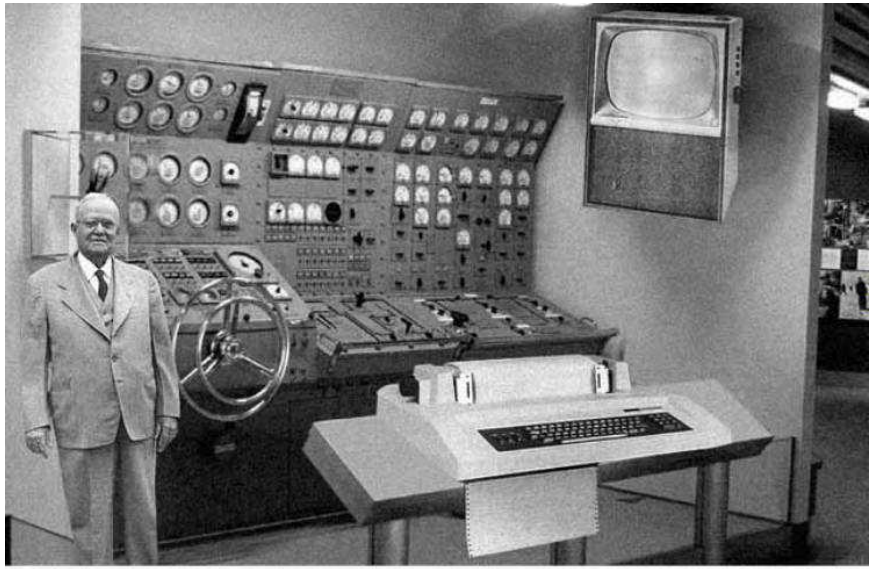
Container-Port Model

- 8 west-coast container ports
  - Punta Colonet, Mexico (proposed, maybe in model)
- Ship attributes: DOT/MARAD
  - No. containers: 3,500-4,000 TEU (Panamax)
  - Costs: operating, support
  - Draft
  - Interarrival time, destination port
- Port attributes
  - Crane speed, yard-storage capacity
- Split load, route to 1-digit ZIP codes (population)
- Scenarios: normal ops, various incidents

Lessons

- Ports not heavily utilized under normal ops
  - Caveat: we don't have export, empties in our model
- LA/LB appears to be the most critical node
  - Invest in better protection, hardening there
- Mexican port could prove critically valuable in TSI
- Future work
  - Refine model for exports, empties
  - Enhance landside transportation
  - Refine rerouting decision rule
  - Evacuate port during a TSI

We Need Better Simulation Software!



Scientists from the RAND Corporation have created this model to illustrate how a "home computer" could look like in the year 2004. However the needed technology will not be economically feasible for the average home. Also the scientists readily admit that the computer will require not yet invented technology to actually work, but 50 years from now scientific progress is expected to solve these problems. With teletype interface and the Fortran language, the computer will be easy to use.

http://www.youtube.com/watch?v=F0GBtKtun7A&feature=player_embedded

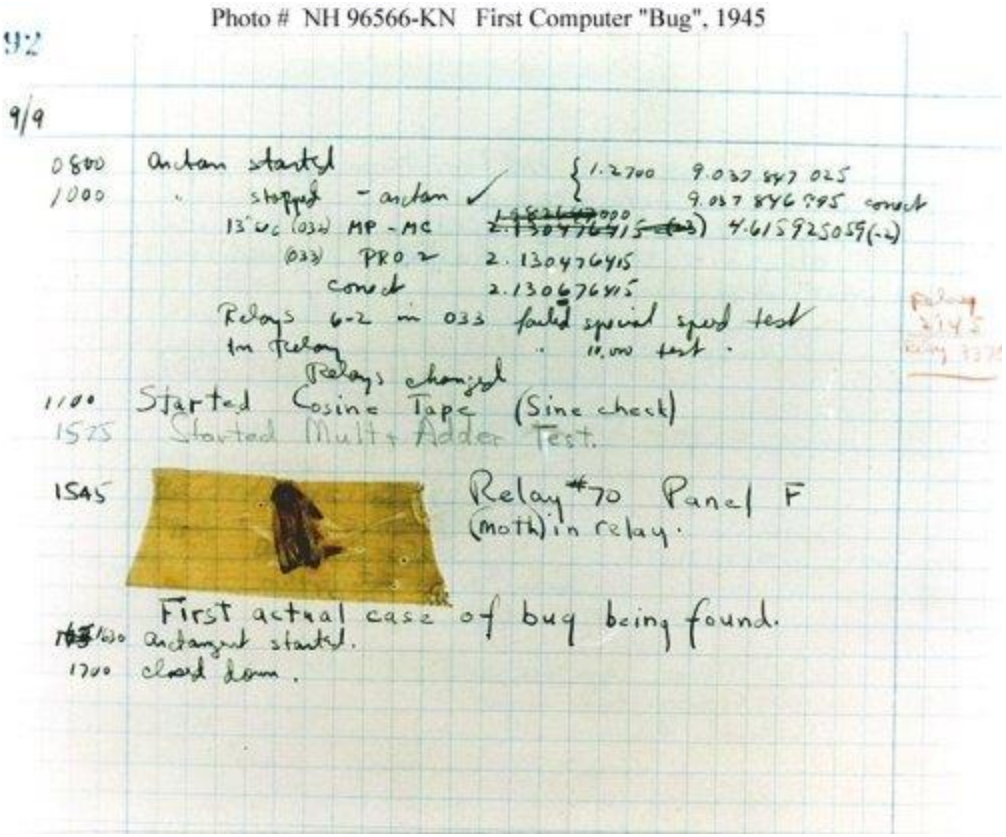
Things We Know How To Do But Aren't (yet universally) in Simulation Software

- Good random-number generators
  - Astronomical cycle length, astronomically many streams and substreams of astronomical length
- Better support for nonstandard input modeling
  - Correlated, multivariate input
  - Nonstationary "arrival" *processes*
  - Autocorrelation within, across input *processes*
  - Empirical, mixed empirical-exponential distributions
  - Integrate real-world database with input modeling

Things We Know How To Do But Aren't (yet universally) in Simulation Software (cont'd.)

- Mindless but correct variance reduction
  - (We already have the mindless part, which everybody does, but it's not correct)
  - Automatic new stream for each source of randomness
- Robust, reliable output analysis
  - Specify precision desired, not run lengths or number of replications
  - Multiple ranking-and-selection methods
  - Optimum seeking with comprehensible stopping conditions

Conclusions



Design and Analysis (Really) Matters

- Thinking carefully about modeling and generating random inputs to simulation has two big advantages
  - Avoid model-validity errors that can occur by going with all the defaults in simulation software (and, as we've seen, these can definitely matter to output)
  - Take advantage of the opportunity to improve precision by smart allocation of random numbers

Practical Simulation Modeling/Analysis

- Understand the system, goals, questions
- Involve people from the system, whether they know (or care) about simulation
- Right level of detail ... more is *not* always better
- Statistical design and analysis ... imperative
  - If you don't have time to do this, don't simulate at all
- Modeling is lots of art, experience ... some science