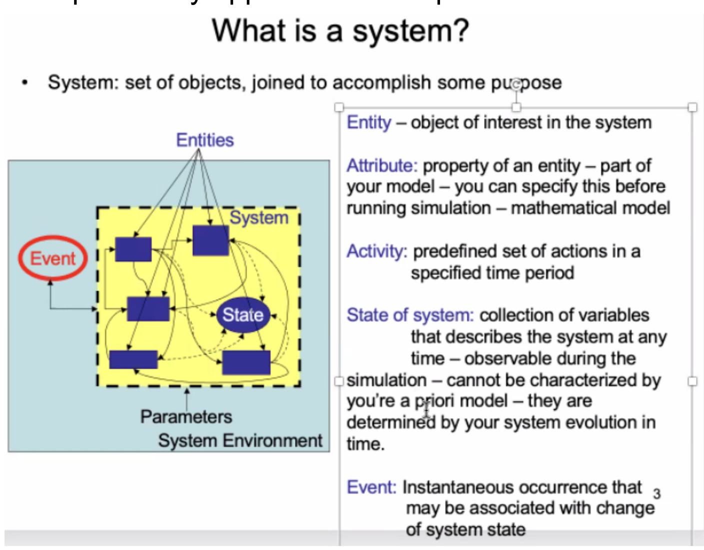
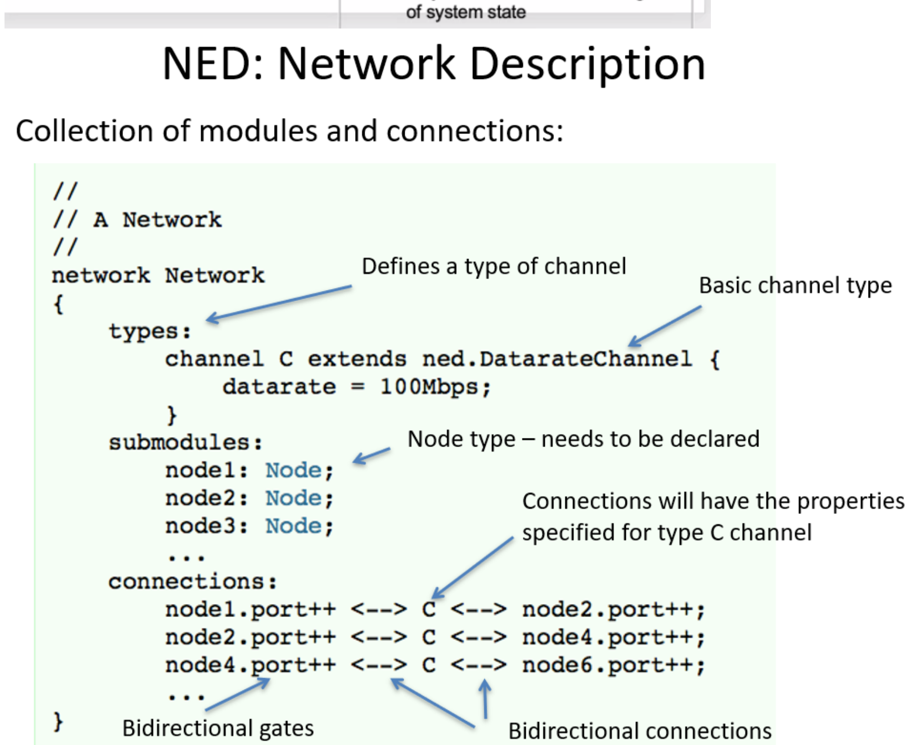
- Know what a discrete event driven dynamic simulation is:
  - **Stochastic**: having a random probability distribution or pattern that may be analyzed statistically but may not be predicted precisely
  - vs **Deterministic**: a system in which no randomness is involved in the development of future states of the system

  - **Continuous:** a system is modeled with the help of variables that change continuously according to a set of differential equations.
  - vs **Discrete:** models the operation of a system as a sequence of events in time, marking a change of state in the system

  - **Dynamic:** model the time-varying behavior by ordinary differential equations or partial differential equations
  - vs **static:** simulation model which has no internal history of both output and input values that were previously applied. It also represents a model in which time is not a factor





- Be able to identify system entities, attributes, actions, events, state variable for a described system
  - Entities (components in the system) like a module
  - Attributes – properties of the entities
  - Actions – how entities respond to events
  - State variable – the collection of variable that completely characterizes a system at a given time (given the objective – measured performance metric)
  - Event – times when state variable changes like a message
- **Time advance algorithm:**The sequence of actions which a simulator (or simulation language) must perform to advance the clock and build a new system snapshot
- **FEL:** Future Event List - those shitty tables in the practice quizzes
- Be able to model stochastic processes based on process description and on the properties and **PMF**s and **PDF**s discussed in class, and compute basic metrics: mean, var, cdf, probabilities
  - Probability Density function: for continuous random variables
  - Probability Mass function: for discrete random variables
- Be able to estimate mean, variance from data and construct **empirical pdf, cdf:** constructed from real data measurements

- **DISCRETE DISTRIBUTIONS:** (I think p(x) is PMF)

- **Bernoulli Trials**: Success or failure
  - For one trial, the Bernoulli distribution is
  - $$p(x)=\begin{cases} p & x=1 \\ 1-p=q & x=0 \\ 0 & ow \end{cases} \qquad \begin{aligned} E(X) &= 0\cdot q+1\cdot p=p \\ \operatorname{var}(X) &= E(X^2)-E(X)^2 = \\ &= \left[0^2\cdot q+1^2\cdot p\right]-p^2=p(1-p) \end{aligned}$$
- **Binomial Distribution** (The number of successes in a Bernoulli process has a binomial distribution) PMF of k successes given n independent events each with a probability p of success
  - $$p(x)=\begin{cases} \binom{n}{x} p^x q^{n-x} & x=1,2,..,n \\ 0 & ow \end{cases} \qquad \begin{aligned} E(X) &= np \\ \operatorname{var}(X) &= npq \end{aligned}$$
- **Geometric distribution:** probability of the number of Bernoulli trials before the first success
  - $$p(x)=\begin{cases} q^{x-1}p & x=1,2,\ldots \\ 0 & ow \end{cases} \qquad E(X)=\frac{1}{p} \qquad \operatorname{var}(X)=\frac{q}{p^2}$$
- **Poisson distribution:** (good for arrival processes) probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event
  - $$p(x)=\begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & x=0,1,2,\ldots \\ 0 & ow \end{cases} \qquad E(X)=\operatorname{var}(X)=\lambda$$

- **CONTINUOUS DISTRIBUTIONS:** (I think f(x) is PDF)

- **Uniform distribution:** has constant probability
  - $$f(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & ow\end{cases}\qquad E(X)=\frac{a+b}{2}\qquad \mathrm{var}(X)=\frac{(b-a)^2}{12}$$
- **Normal distribution (Gaussian distribution):** standard bell curve

  Mean μ, variance σ^2

  Mode and mean are equal
  - $$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right)}\qquad F(x)=P(X\le x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$$
- **Rayleigh distribution:** fast fading
  - $$f(x)=\frac{x}{p}\exp\left(-\frac{x^2}{2p}\right), \mathrm{x}\ge 0\quad E(X)=\sqrt{\frac{\pi}{2}p}\,;\quad \mathrm{var}(X)=\frac{4-\pi}{2}\mathrm{p}$$
- **Lognormal distribution:** (If X is lognormal, ln(X) is normally distributed with mean mu and variance sigma^2)
  - pdf: $$f(x)=\frac{1}{\sqrt{2\pi}\sigma x}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right],\quad \mathrm{x}>0\qquad \mu_L=e^{\mu+\sigma^2/2}\qquad \sigma_L^2=e^{\sigma^2+2\mu}\left(e^{\sigma^2}-1\right)$$
- **Exponential distribution:** inter-arrival times and service times
  - $$E(X)=\frac{1}{\lambda}\qquad \mathrm{var}(X)=\frac{1}{\lambda^2}$$ *pdf:* $$f(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & ow\end{cases}$$

  - False alarm probability: Miss-detection probability:
- $$P_{\epsilon_0}=\int_{\mu}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-y^2/(2\sigma^2)}dy.\qquad P_{\epsilon_1}=\int_{-\infty}^{\mu}\frac{1}{\sqrt{2\pi}\sigma}e^{-(y-A)^2/(2\sigma^2)}dy$$
- **Erlang Distribution:** when you have multiple exponentials in series
  - If you add *k* independent exponential random variables, with rate λ, the resulting random variable has an Erlang distribution of order *k:*

    $$f(x)=\frac{\lambda^k e^{-\lambda x}x^{k-1}}{(k-1)!},\qquad \mathrm{x}\ge 0$$
  - *For k=1 → exponential*
  - *CDF:*

    $$F(x)=1-\sum_{i=0}^{k-1}\frac{(\lambda x)^i e^{-\lambda x}}{i!}$$
  - *Mean and variance:*

    $$E(X)=\frac{k}{\lambda};\quad \mathrm{var}(X)=\frac{k}{\lambda^2}$$

- **Quantile-Quantile plots**: probability plot comparing two probability distributions by plotting their quantiles against each other – near y=x if distributions are similar. Linear related will form a line (not necessarily y=x)
- Understand **Chi-Square test** and be able to apply for given data
    - to determine whether there is a statistically significant difference between the expected frequencies and the observed frequencies
    - Break into classes (bins) - count number of samples in each class O_i - compute expected number of samples in each if it were Gaussian E_i - compute error: $\frac{(O_i - E_i)^2}{E_i}$ - add error across classes - compare with threshold (like 0.01) - reject H_0 if error > threshold
    - K-s-1 degrees of freedom. s is number of parameters estimated
- Dependence: Covariance and Correlation:
    - If X_1 and X_2 are two r.v. with mean μ_i and variance σ_i^2

      Covariance:

      $$\text{cov}(X_1, X_2) = E[(X - \mu_1)(X - \mu_2)] = E(X_1 X_2) - \mu_1 \mu_2$$

      Properties: $-1 \le \rho \le 1$

      – Correlated if $|\rho| \ge 0.33$

      – Uncorrelated if $|\rho| \le 0.33$

      Correlation:

      $$\rho = corr(X_1, X_2) = \frac{\text{cov}(X_1 X_2)}{\sigma_1 \sigma_2}$$
- Understand components of queuing systems and identify type of queues (based on the
- **A/B/c/K/N** notation
    - A: Type of arrival in the queues
        - M is exponentially or poisson distributed
        - D is constant/deterministic
        - G is arbitrary/general
    - B is service time distribution - same letters as above
    - c is number of servers in parallel
    - N is capacity of the queue (number in queue + number in service)
    - K number of population (finite or infinite) – often dropped from notation if infinite
- Queueing notation for parallel server systems

P_n	Steady-state probability of having *n* customers in the system
$P_n(t)$	Probability of *n* customers in system at time t
λ	Arrival rate
λ_e	Effective arrival rate
μ	Service rate of one server
ρ	Server utilization
A_n	Inter-arrival time between customer *n-1* and *n*
S_n	Service time of the n^{th} arriving customer
W_n	Total time spent in the system by the n^{th} arriving customer
W_n^Q	Total time spent in the waiting line by the n^{th} arriving customer
$L(t)$	Number of customers in system at time t
$L_Q(t)$	The number of customers in queue at time t
L	Long-run time-average number of customers in the system
L_Q	Long-run time-average number of customers in queue
w	Long-run average time spent in system per customer
w_Q	Long-run average time spent in queue per customer

- **Queue stability condition**: service time < arrival time (arrival rate < service rate) **CHECK THIS**
- Be able to solve for time in system, time in queue, number of customer in queue, number of customers in system, server utilization
  - ??? formulas?

  The conservation equation (Little equation)

  - Little's equation: $\boxed{L = \lambda w}$ as $T \to \infty$, $N \to \infty$
-
- *^The average number of customers in the system at an arbitrary point in time is equal to the average number of arrivals per unit time, times the average time spent in the system.*
- **Server Utilization:** percentage of time that server is busy (serving customers) ρ
  - For a general queue: G/G/1 (λ = arrival rate, μ = service rate) $\rho = \frac{\lambda}{\mu}$
  - And stability condition $\lambda < \mu \Rightarrow \rho < 1$
  - For **G/G/c** $\rho = \frac{\lambda}{c\mu}$ and stability condition $\lambda < c\mu \Rightarrow \rho < 1$

  - Average number of customers in the system

$$L = \sum_{n=0}^{\infty} nP_n$$

  - Average customer time in the system

$$w = \frac{L}{\lambda} \quad \text{(we have used Little's equation)}$$

  - Average customer time in queue

$$w_Q = w - \frac{1}{\mu} \quad \text{(time in system – service time)}$$

  - Average number of customers in queue

$$L_Q = \lambda w_Q \quad \text{(again Little's equation)}$$

Steady-state formulas for M/G/1

Mean service time $1/\mu$, service variance σ^2

ρ	λ / μ
L	$\rho + \frac{\lambda^2(1/\mu^2+\sigma^2)}{2(1-\rho)} = \rho + \frac{\rho^2(1+\sigma^2\mu^2)}{2(1-\rho)}$
w	$\frac{1}{\mu} + \frac{\lambda(1/\mu^2+\sigma^2)}{2(1-\rho)}$
w_Q	$\frac{\lambda(1/\mu^2+\sigma^2)}{2(1-\rho)}$
L_Q	$\frac{\lambda^2(1/\mu^2+\sigma^2)}{2(1-\rho)} = \frac{\rho^2(1+\sigma^2\mu^2)}{2(1-\rho)}$
P_0	$(1-\rho)$

-

M/M/c parameters

ρ	$\lambda / c\mu$
P_0	$\left\{\left[\sum_{n=0}^{c-1}\frac{(\lambda/\mu)^n}{n!}\right]+\left[\left(\frac{\lambda}{\mu}\right)^c\left(\frac{1}{c!}\right)\left(\frac{c\mu}{c\mu-\lambda}\right)\right]\right\}^{-1} = \left\{\left[\sum_{n=0}^{c-1}\frac{(c\rho)^n}{n!}\right]+\left[(c\rho)^c\left(\frac{1}{c!}\right)\left(\frac{1}{1-\rho}\right)\right]\right\}^{-1}$
$P(L(\infty) \geq c)$	$\frac{(\lambda/\mu)^c P_0}{c!(1-\lambda/c\mu)} = \frac{(c\rho)^c P_0}{c!(1-\rho)}$
L	$c\rho + \frac{(c\rho)^{c+1}P_0}{c c!(1-\rho)^2} = c\rho + \frac{\rho P(L(\infty)\geq c)}{(1-\rho)}$
w	L/λ
w_Q	$w - 1/\mu$
L_Q	λw_Q
$L-L_Q$	$\lambda/\mu = c\rho$

20

Steady-state parameters for M/G/∞ queue

P_0	$e^{-\lambda/\mu}$
w	$1/\mu$
w_Q	0
L	λ/μ
L_Q	0
P_n	$\frac{e^{-\lambda/\mu}(\lambda/\mu)^n}{n!}, \quad n = 0,1,2,\ldots$

M/M/1 queue

Service times – also exponential, with mean $1/\mu$

– From exponential distribution: $\sigma^2 = 1/\mu^2$

L	$\frac{\lambda}{\mu-\lambda} = \frac{\rho}{1-\rho}$
w	$\frac{1}{\mu-\lambda} = \frac{1}{\mu(1-\rho)}$
w_Q	$\frac{\lambda}{\mu(\mu-\lambda)} = \frac{\rho}{\mu(1-\rho)}$
L_Q	$\frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{\rho^2}{\mu(1-\rho)}$
P_n	$\left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n = (1-\rho)\rho^n$

Steady state parameters for the M/M/c/N queue:

$a = \lambda/\mu,\ \rho = \lambda/c\mu$

P_0	$\left[1+\sum_{n=1}^{c}\frac{a^n}{n!}+\frac{a^c}{c!}\sum_{n=c+1}^{N}\rho^{n-c}\right]^{-1}$
P_N	$\frac{a^N}{c!c^{N-c}}P_0$
L_Q	$\frac{P_0 a^c \rho}{c!(1-\rho)^2}\left[1-\rho^{N-c}-(N-c)\rho^{N-c}(1-\rho)\right]$
λ_e	$\lambda(1-P_N)$
w_Q	L_Q/λ_e
w	$w_Q + 1/\mu$
L	$\lambda_e w$

Also, for infinite calling population, we must make sure that the system is stable:

- $$\rho = \frac{\lambda}{c\mu} < 1$$
- M/G/1 - For L_Q and W_Q, correction for L_Q and W_Q: multiply M/M/1 formulas with $\frac{1+(cv)^2}{2}$
- M/G/c - no exact formula, approximate with M/M/c multiplied by same correction factor

Time average number in queue: L_Q

– Same reasoning as before, leads to

$$\hat{L}_Q = \frac{1}{T}\int_0^T L_Q(t)dt \rightarrow L_Q, \text{as } T \rightarrow \infty$$

– Average time spent in system by customer

- N = number of arrivals during [0,T]

$$\hat{w} = \frac{1}{N}\sum_{i=0}^{\infty} W_i \rightarrow w, \text{as } T \rightarrow \infty, \text{N} \rightarrow \infty$$

Average time spent in queue by customer

$$\hat{w}_Q = \frac{1}{N}\sum_{i=0}^{\infty} W_i^Q \rightarrow w_Q, \text{as } T \rightarrow \infty,\ \text{N} \rightarrow \infty$$

- Ti = total time during [0,T], in which the system contained exactly i customers
- $$\hat{L} = \frac{1}{T}\sum_{i=0}^{\infty} iT_i = \sum_{i=0}^{\infty} i\left(\frac{T_i}{T}\right)$$
- **Queue disciplines**
  - FIFO: first-in-first-out
  - LIFO: last-in-first-out
  - SIRO: service in random order
  - SPT: shortest processing time first
  - PR: service according to priority