Basic Statistics

- Conditional Prob: $P(A|B) = P(A \cap B)/P(B)$
- Independent $\iff P(A \cap B) = P(A)P(B)$
- $Var[X] = E[X^2] - E^2[X]$
- $E[aX+b] = aE[X]+b$, $V[aX+b] = a^2V[X]$
- Joint CDF: $F(x,y) \equiv P(X \le x, Y \le y),\ \forall x, y$
- Joint PDF: $f(x,y) \equiv \frac{\partial^2}{\partial x \partial y} F(x,y)$
- Marginal CDF: $F_X(x) = \int F(x,y)dy$
- Conditional PDF: $f(y|x) \equiv f(x,y)/f_X(x)$
- $E[XY] = \iint xyf(x,y)dxdy$
- $Cov(X,Y) = E[XY] - E[X]E[Y]$
  $Cov(aX, bY) = abCov(X,Y)$
  $Cov(X \pm Y, Z) = Cov(X,Z) \pm Cov(Y,Z)$
- $Corr(X,Y) \equiv Cov(X,Y)/\sqrt{V(X)V(Y)}$
  $Corr(aX, bY) = Corr(X,Y)$
- Independent $\Rightarrow Cov(X,Y) = 0$
- $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X,Y)$
- Bayes' Formula:

$$\Pr(E_i|O_j) = \frac{\Pr(O_j|E_i)\Pr(E_i)}{\sum \Pr(O_j|E_i)\Pr(E_i)}$$

Probability Distributions

- $X \sim \text{Bernoulli}(p)$:

$$f(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$$

  $E[X] = p$, $Var(X) = p(1-p)$

- $X \sim \text{Binomial}(n,p)$:
  (# of successes in n Bern(p) trials)

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

  $E[X] = np$, $Var(X) = np(1-p)$

- $X \sim \text{Geometric}(p)$:
  (# of Bern(p) trials until a success occurs)

$$f(x) = (1-p)^{x-1}p$$

  $E[X] = 1/p$, $Var(X) = (1-p)/p^2$

- $X \sim \text{Poisson}(\lambda)$:

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!},\ x = 0, 1, \ldots$$

  $E[x] = \lambda = Var(X)$

- $X \sim \text{Uniform}(a,b)$:

$$f(x) = \frac{1}{b-a},\ E[X] = \frac{a+b}{2},\ V(X) = \frac{(b-a)^2}{12}$$

- $X \sim \text{Exponential}(\lambda)$: time between Poi events

$$f(x) = \lambda e^{-\lambda x},\ F(x) = 1 - e^{-\lambda x},$$

  $E[X] = 1/\lambda$, $Var(X) = 1/\lambda^2$. And

$$P(X > s+t | X > s) = P(X > t)$$

- $X \sim \text{Erlang}(k, \lambda)$: the sum of k Exp(λ)

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!},\ for\ x, \lambda \ge 0$$

$$F(x) = 1 - \sum_{n=0}^{k-1} \frac{e^{-\lambda x}(\lambda x)^n}{n!}$$

  $E(X) = k/\lambda$, $Var(X) = k/\lambda^2$

- $X \sim \text{Triangular}(a,b,c)$: $E(X) = (a+b+c)/3$
- $X \sim \text{Normal}(\mu, \sigma^2)$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],\ x \in \mathbb{R}$$

  $E[X] = \mu$, $V[X] = \sigma^2$, $M_X(t) = exp[\mu t + \frac{1}{2}\sigma^2 t^2]$

- *Law of Large Numbers* (special case):

$$X_1, \ldots, X_n \text{ are iid Nor} \Rightarrow \bar{X}_n \sim \text{Nor}(\mu, \sigma^2/n)$$

- *Central Limit Theorem*: If $X_1, \ldots, X_n \overset{iid}{\sim} f(x)$,

$$Z_n \equiv \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} \text{Nor}(0,1)$$

Input Modeling

Represent the uncertainty in a stochastic simulation.

- Fundamental Requirements: 1. capable of representing the physical realities of the process; 2. easily tuned to the situation on hand; and 3. amenable to random variate generation.

- Input Modeling with Data: 1. select one or more candidate distribution, based on physical characteristics of the process and graphical examination of the data; 2. fit the distribution with data; 3. check the fit to the data via tests and graphical analysis; and 4. if the distribution does not fit, select another candidate and go to 2, or use an empirical distribution.

- Physical Basis for Distributions: 1. **Binomial**: # of successes in n iid Bernoulli(p) trials; 2. **Negative Binomial**: # of trials required to achieve k "successes" (the sum of k iid Geom(p)); 3. **Poisson**: # of independent events that occur in a fixed amount of time or space; 4. **Normal**: the distribution of a process that can be thought of as the sum of a number of component processes; 5. **Lognormal**: the distribution of a process that can be thought of as the product of a number of component processes; 6. **Exponential**: the time between independent events, or a process time which is memoryless; 7. **Erlang**: the sum of k identical exponential random variables; 8. **Gamma**: an extremely flexible distribution used to model nonnegative random variables; 9. **Beta**: an extremely flexible distribution used to model bounded (fixed upper and lower limits) random variables; 10. **Weibull**: the time to failure for components, can model increasing or decreasing failure rate hazard; 11. **Uniform**: models complete uncertainty, since all outcomes are equally likely; 12. **Triangular**: models a process when only the minimum, most likely and maximum values of the distribution are known; 13. **Empirical**: reuses the data themselves by making each observed value equally likely, can be interpolated to obtain a continuous distribution.

- Common Methods for Fitting: maximum likelihood, method of moments, and least squares. (While the method matters, the variability in the data often overwhelms the differences in the estimators.)

- Ways to Check Fit: χ^2, K-S, Anderson-Darling tests; density-histogram, and q-q plots.

- p-value: Type I error level (significance) at which we would just reject H_0 for the given data. (less likely to reject H_0 at larger p-value)

- q-q Plot: displays the sorted data ($Y_1 \le Y_2 \le \cdots \le Y_n$) vs $F^{-1}((j-1/2)/n), j = 1, 2, \ldots, n$. **Features**: 1. does not depend on how the data are grouped; 2. better than density-histogram when the number of data points is small; and 3. deviations from a straight line show where the distribution does not match (a straight line implies the family of distributions is correct; a 45° line implies correct parameters, a curved line implies a wrong dist'n family).

- χ^2 Test: 1. a formal comparison of a histogram or line graph with the fitted density or mass function; and 2. sensitive to how we group the data.

- K-S and A-D Test: 1. comparison of an empirical distribution function with the distribution function of the hypothesized distribution; 2. does not depending on the grouping of data; and 3. A-D detects discrepancies in the tails and higher power than K-S test.

- Beware of goodness-of-fit tests because they are unlikely to reject any distribution when you have little data, and are likely to reject every distribution when you have lots of data; Avoid histogram-based summary measures, if possible, when asking the software for its recommendation.

- Empirical Distribution: 1. As the sample size goes to infinity, the empirical distribution converges to "the truth"; 2. no assumed distribution need to be selected; and 3. only the values we saw can appear again, no tails and nothing in the gaps.

- Interpolated Empirical: to fill in gaps, we linearly interpolate between the sorted data points.

- Breakpoints Method: useful for modeling quantities with a large number of possible outcomes. (smallest and largest possible values, most likely value, 1-3 breakpoints)

- Mean & Variability Method: ..., also useful for modeling the variability in percentage changes.

(mean value, an average percentage variation around that mean, upper and lower limits)

- Correlations: if data exists, calculate sample correlation; if not, percentage of the time the two inputs move together ($P\%$), then $|\rho| = P/100$.

Inventory Management

- News Vendor Problem: total cost is
$$c_p \min\{D, y\} + c_s(y-D)^+ - c_v y \Rightarrow$$
$$c_p D - \{c_v y + c_p(D-y)^+ - c_s(y-D)^+\}$$
optimal quantity: $F(y^*) \geq (c_p - c_v)/(c_p - c_s)$

- EOQ with Certain Demand:
$$\mathrm{TC}(q) = \frac{KD}{q} + pD + \frac{hq}{2} \Rightarrow q^* = \sqrt{\frac{2KD}{h}}$$

- EOQ with Uncertain Demand (back-ordered):
$$K\frac{\mathrm{E}[D]}{q} + h(r - \mathrm{E}[X] + \frac{q}{2}) + c_B \mathrm{E}[B]\frac{\mathrm{E}[D]}{q}$$
$$q^* = \sqrt{\frac{2\mathrm{E}[D](K + c_B \mathrm{E}[B])}{h}}$$
$$\Pr(X \geq r^*) = \frac{hq^*}{c_B \mathrm{E}[D]}$$

- # of Backorder: $\mathrm{E}[B] = \sigma_X L_{SN}(\frac{r-\mu_X}{\sigma_X})$

- Safety Stock: $ss = r - \mathrm{E}[X]$

- Ordering & Transportation: $p\mathrm{E}[D] + K\mathrm{E}[D]/q$
Pipeline Inventory: $\mathrm{E}[D]ivL$
Facility Inventory: $iv(q/2 + ss) + s(q + ss)$
Backordered demand: $c_B \mathrm{E}[B]\mathrm{E}[D]/q$

- EOQ with Uncertain Demand (loss sales):
$$K\frac{\mathrm{E}[D]}{q} + h(r - \mathrm{E}[X] + \frac{q}{2} + \mathrm{E}[B]) + c_{LS}\mathrm{E}[B]\frac{\mathrm{E}[D]}{q}$$
$$q^* = \sqrt{\frac{2\mathrm{E}[D](K + c_{LS}\mathrm{E}[B])}{h}}$$
$$\Pr(X \geq r^*) = \frac{hq^*}{hq^* + c_{LS}\mathrm{E}[D]}$$

- Service Level Measure 1: fraction of all demand D that is met on time ($\beta = 1 - \mathrm{E}[B]/q$).

- Service Level Measure 2: proportion of cycles in which no shortage occurs ($\alpha = 1 - \Pr(X > r)$).

- Standard Normal Loss Function:
$$L_{SN}(z) = \varphi(z) - z[1 - \Phi(z)], \varphi(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}$$

- L is constant: $D_{R+L} \sim \mathrm{Nor}\{(R+L)\mathrm{E}[D], (R+L)\mathrm{V}[D]\}$; otherwise: $D_{R+L} \sim \mathrm{Nor}\{R\mathrm{E}[D] + \mathrm{E}[L]\mathrm{E}[D], R\mathrm{V}[D] + \mathrm{E}[L]\mathrm{V}[D] + \mathrm{E}^2[D]\mathrm{V}[L]\}$

- Order Up to Policy (back-ordered):
$$p\mathrm{E}[D] + \frac{K+J}{R} + h(S - \mathrm{E}[D_{R+L}] + \frac{1}{2}\mathrm{E}[D_R]) + c_B\frac{\mathrm{E}[B]}{R}$$
$$\mathrm{EOQ} = \sqrt{\frac{2(K+J)\mathrm{E}[D]}{h}}$$
$$\Pr(D_{R+L} \geq S) = hR/c_B$$

- $\mathrm{E}[B] = \sqrt{(R+L)\mathrm{V}[D]} L_{SN}\left(\frac{S - (R+L)\mathrm{E}[D]}{\sqrt{(R+L)\mathrm{V}[D]}}\right)$

- Find S that $\Pr(D_{R+L} > S) = 1 - \alpha$:
$$S = (R+L)\mathrm{E}[D] + z_{1-\alpha}\sqrt{(R+L)\mathrm{V}[D]}$$

- Fill Rate: $\mathrm{E}[B]/(R\mathrm{E}[D]) = 1 - \beta$

- Order Up to Policy (loss sales):
$$\Pr(D_{R+L} \geq S) = hR/(hR + c_{LS})$$

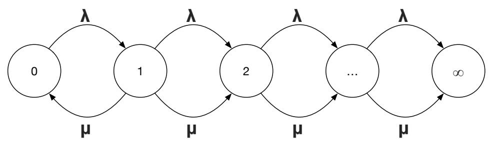
- Ordering & Transportation: $p\mathrm{E}[D] + K/R$
Pipeline Inventory: $\mathrm{E}[D]ivL$
Facility Inventory: $h(\frac{R\mathrm{E}[D]}{2} + ss) + s(R\mathrm{E}[D] + ss)$
Back-ordered demand: $c_B \mathrm{E}[B]/R$

- **Annually Cost $\Rightarrow$ Weekly!**

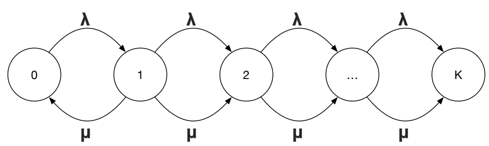
Queuing System (A/B/C/K/N)

- A: interarrival time λ (M/EXPO, D/DETER, E/ERLANF, G/GENERAL); B: service time μ; C: # of identical and parallel servers; K: system capacity (buffer + servers); N: population size

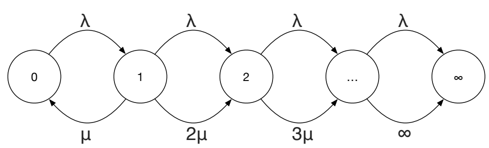
- $X(t)$: # of customer in system at time t;
$\pi = \{\pi_0, \pi_1, \dots\}$: probability of n in queue;
$\rho = 1 - \pi_0$: fraction of time that a server is busy;
$L = \sum_{i=0}^{\infty} i\pi_i$: # of customer in the system;
$L_q = \sum_{i=1}^{\infty}(i-1)\pi_i$: # of customer queuing;
$W = L/\lambda_{\mathrm{eff}}$: average time in system;
$W_q = L_q/\lambda_{\mathrm{eff}}$: average time queuing



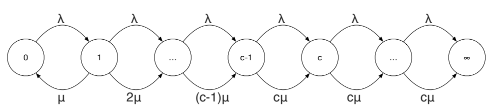
- M/M/1: (little's law: $L = \lambda_{eff} \times W$)
  $\pi_n = (\lambda/\mu)^n \pi_0$, $\pi_0 = (\lambda/\mu)^n(1 - \lambda/\mu)$
  $\rho = 1 - \pi_0 = \lambda/\mu$
  $L = \lambda/(\mu - \lambda)$, $L_q = \lambda^2/(\mu^2 - \mu\lambda)$
  $W = 1/(\mu - \lambda)$, $W_q = \lambda/(\mu^2 - \mu\lambda)$



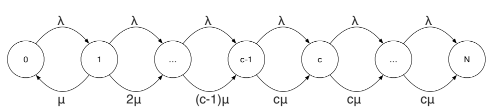
- M/M/1/K: $\mu\pi_1 = \lambda\pi_0$, $\mu\pi_k = \lambda\pi_{k-1}$
  $\lambda\pi_{i-1} + \mu\pi_{i+1} = (\mu + \lambda)\pi_i, \ (1 \le i \le k-1)$



- M/M/∞: $\pi_0 = e^{-\lambda/\mu}$, $\pi_n = e^{-\lambda/\mu}(\lambda/\mu)^n/n!$



- M/M/C: $\rho = \lambda/(c\mu)$, $L_q = \sum_{i=1}^{\infty} i\pi_{c+i}$
  $W_q = L_q/\lambda$, $W = W_q + 1/\mu$



- M/M/C/N: $\lambda_{\text{eff}} = \lambda \Pr(\text{accept})$
- G/G/1: for large $\rho = \lambda/\mu$, $W_q = \frac{1}{\mu}\frac{\rho}{1-\rho}\frac{c_a^2 + c_s^2}{2}$
  $c_a^2 = \text{Var}(\text{interarrival time})/\text{E}^2(\text{interarrival time})$
  $c_s^2 = \text{Var}(\text{service time})/\text{E}^2(\text{service time})$
- G/G/C: $W_q = (W_q \text{ for M/M/C}) \times (1 + c_s^2)/2$