General Intro / Overview
Simulation Scope of this class: Discrete, Stochastic, and Dynamic
Simulating Pi = Area of Unit Circle / Area of Unit Square = pi/4
Simulating Integral: Sample n rectangle on [0,1], with width 1/n and height f(x) – centered randomly on [0,1]. Add up to approximate integral of f(x)
Random Walk: Converges to Brownian Motion
Generate Unif(0,1) pseudo random numbers (PRNS): deterministic algo

- Integer Seed X(0) - X(i) = a*X(i-1)mod(m). ith PRN U(i) = X(i)/m
- **Common**: 16807 X(i-1) mod(2^{31}-1)

Generate RVs – start with Unif(0,1), apply inverse transform to get RV

- Example: -(1/λ) ln(U(i)) ~ Exp(λ)

Output analysis: Independent Replications are sampled, means, then CLT. Traditional statistics on the sample means, not original samples.

- Steady State: Must be warmed up, then batch up long running data and calculate sample means from each batch

Calculus Overview

$\frac{d}{dx}f(x) \equiv f'(x) \equiv \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$

Common Deriv and Properties

$[\sin(x)]' = \cos(x), [\cos(x)]' = -\sin(x), [\ell n(x)]' = 1/x,$
$[\arctan(x)]' = 1/(1+x^2), [e^x]' = e^x,$

$$[af(x)+b]' = af'(x), \quad [f(x)+g(x)]' = f'(x)+g'(x),$$

$$[f(x)g(x)]' = f'(x)g(x)+f(x)g'(x) \quad \text{(product rule)},$$

$$\left[\frac{f(x)}{g(x)}\right]' = \frac{g(x)f'(x)-f(x)g'(x)}{g^2(x)} \quad \text{(quotient rule)}^1,$$

$$[f(g(x))]' = f'(g(x))g'(x) \quad \text{(chain rule)}^2.$$

Minimum / Maximum can only occur where derivative = 0. If f''(x) < 0 it's a max, if it's >0 it's a min. If it's = 0 it's a point of inflection.

Solving for Zero: Trial and Error, Bisection, Newton's, Fixed Point

Bisection: Find X1, X2 such that g(X1) < 0 and g(X2) > 0. X3 = (X1+X2)/2. If g(x3) < 0, the 0 is in [X2, X3] and vice versa

Newtons:

$$x_{i+1} = x_i - \frac{g(x_i)}{g'(x_i)}$$

Integration

$$\int_a^b f(x)\,dx \equiv F(x)\Big|_a^b \equiv F(b)-F(a)$$

$$\int x^k\,dx = \frac{x^{k+1}}{k+1}+C,\ k\neq -1 \quad \int\frac{dx}{x} = \ell n|x|+C,$$

$$\int e^x\,dx = e^x+C, \quad \int\cos(x)\,dx = \sin(x)+C,$$

$$\int\frac{dx}{1+x^2} = \arctan(x)+C.$$

$$\int_a^a f(x)\,dx = 0,$$

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx,$$

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx+\int_c^b f(x)\,dx.$$

$$\int[f(x)+g(x)]\,dx = \int f(x)\,dx+\int g(x)\,dx,$$

$$\int f(x)g'(x)\,dx = f(x)g(x)-\int g(x)f'(x)\,dx \quad (\text{parts})^4,$$

$$\int f(g(x))g'(x)\,dx = \int f(u)\,du \qquad \sum_{k=1}^n k = \frac{n(n+1)}{2},$$

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}, \quad \sum_{k=0}^\infty p^k = \frac{1}{1-p} \quad (\text{for } -1<p<1).$$

L'Hopital:

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}.$$

Methods for Approximating an Integral when no Closed Form solution:

Reimann:

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^n f(x_i)\Delta x = \frac{b-a}{n}\sum_{i=1}^n f\left(a+\frac{i(b-a)}{n}\right)$$

Trapezoid Integr:

$$\frac{b-a}{n}\left[\frac{f(a)}{2}+\sum_{i=1}^{n-1}f\left(a+\frac{i(b-a)}{n}\right)+\frac{f(b)}{2}\right].$$

Monte Carlo Integr:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{n}\sum_{i=1}^n f(a+(b-a)U_i),$$

Probability Overview

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

If Numerator = P(A)P(B), then independent, and P(A|B) = P(A). PMF = Discrete, PDF = Continuous. Both sum up to 1.

$$\textstyle\sum_x f(x) = 1. \quad \int_{\mathbb{R}} f(x)\,dx = 1.$$

$$F(x) \equiv P(X\le x) = \begin{cases}\sum_{y\le x} f(y) & \text{if } X \text{ is discrete}\\ \int_{-\infty}^x f(y)\,dy & \text{if } X \text{ is continuous}\end{cases}$$

Inverse Transform Theorem: if X is a continuous RV with cdf F(x), then random variable F(X) ~ U(0,1). Thus we can generate X by setting X = F^{-1}(U)

$$\mathrm{E}[X] \equiv \begin{cases}\sum_x xf(x) & \text{if } X \text{ is discrete}\\ \int_{\mathbb{R}} xf(x)\,dx & \text{if } X \text{ is continuous}\end{cases} = \int_{\mathbb{R}} x\,dF(x)$$

LOTUS, where h(X) is some function of the RV X. h(X) can be 1/X, ln(X), etc.

$$\mathrm{E}[h(X)] = \begin{cases}\sum_x h(x)f(x) & \text{if } X \text{ is disc}\\ \int_{\mathbb{R}} h(x)f(x)\,dx & \text{if } X \text{ is cts}\end{cases} = \int_{\mathbb{R}} h(x)\,dF(x).$$

$\mathrm{Var}(X) = \mathrm{E}[X^2]-(\mathrm{E}[X])^2$. StDev is sqrt(Var(X))

$$\mathrm{E}[aX+b] = a\mathrm{E}[X]+b \text{ and } \mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$$

Definition: $M_X(t) \equiv \mathrm{E}[e^{tX}]$ is the *moment generating function* (mgf) of the RV X. ($M_X(t)$ is a function of t, *not* of X!)

$$\mathrm{E}[X^k] = \frac{d^k}{dt^k}M_X(t)\Big|_{t=0}, \quad k = 1,2,\ldots.$$

. Useful for E[X^2]

$$\begin{aligned}P(Y\le y) &= P(F(X)\le y)\\ &= P(X\le F^{-1}(y))\\ &= F(F^{-1}(y)) = y,\end{aligned}$$

$$\begin{aligned}G(y) &= P(Y\le y)\\ &= P(X^2\le y)\\ &= P(-\sqrt{y}\le X\le\sqrt{y})\\ &= \int_{-\sqrt{y}}^{\sqrt{y}}|x|\,dx = y, \quad 0<y<1.\end{aligned}$$

The CDF of Y:

$$F_Y(y) = P(Y\le y) = P(h(X)\le y) = P(X\le h^{-1}(y)).$$

Thus the PDF of y:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = f_X(h^{-1}(y))\left|\frac{d}{dy}h^{-1}(y)\right|.$$

Joint PDFs / CDFs

Joint CDF:

$$F(x,y) \equiv P(X\le x, Y\le y)$$

Marginal PDFs:

$$f_X(x) = P(X=x) = \sum_y f(x,y). \quad f_Y(y) = P(Y=y) = \sum_x f(x,y)$$

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy \quad\text{and}\quad f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx.$$

NOTE: when doing dy (fX(x)), you integrate over the range of y values.

X and Y are **independent** if $f(x,y) = f_X(x)f_Y(y)$ for all x,y., and their domains don't depend on one another. EX: c/(x+y) NOT independent.

$$f(y|x) = \frac{f(x,y)}{f_X(x)} \qquad f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)}.$$

Conditional Expectation

$$\mathrm{E}[Y|X=x] \equiv \begin{cases}\sum_y yf(y|x) & \text{discrete}\\ \int_{\mathbb{R}} yf(y|x)\,dy & \text{continuous}\end{cases}$$

Double Expectation

$$\begin{aligned}\mathrm{E}[\mathrm{E}(Y|X)] &= \int_{\mathbb{R}}\mathrm{E}(Y|x)f_X(x)\,dx\\ &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} yf(y|x)\,dy\right)f_X(x)\,dx\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} yf(y|x)f_X(x)\,dx\,dy\\ &= \int_{\mathbb{R}} y\int_{\mathbb{R}} f(x,y)\,dx\,dy\\ &= \int_{\mathbb{R}} yf_Y(y)\,dy = \mathrm{E}[Y]. \quad\Box\end{aligned}$$

$$P(Y<X) = \int_{\mathbb{R}} P(Y<x)f_X(x)\,dx.$$

Example: If $X\sim\mathrm{Exp}(\mu)$ and $Y\sim\mathrm{Exp}(\lambda)$ are indep RV's, then

$$\begin{aligned}P(Y<X) &= \int_{\mathbb{R}} P(Y<x)f_X(x)\,dx\\ &= \int_0^\infty(1-e^{-\lambda x})\mu e^{-\mu x}\,dx\\ &= \frac{\lambda}{\lambda+\mu}. \quad\Box\end{aligned}$$

Covariance and Correlation

Var(X_bar) = Var(X_I)/n. Decreases as N increases

E[X_bar] = E[Xi]

$$\rho \equiv \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)\\ \mathrm{Var}(X-Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)-2\mathrm{Cov}(X,Y).\end{aligned}$$

$$\mathrm{Cov}(aX,bY) = ab\mathrm{Cov}(X,Y).$$

Cov(X, Y) = E{XY] – E[X]E[Y]. Cov = 0 when independent.

Probability Distributions

Bern: $\mathrm{E}[X]=p$, $\mathrm{Var}(X)=pq$, $M_X(t)=pe^t+q$.

Binom: $f(y) = \binom{n}{y}p^y q^{n-y}$.

$\mathrm{E}[Y]=np$, $\mathrm{Var}(Y)=npq$, $M_Y(t)=(pe^t+q)^n$

Geom: $f(x) = q^{x-1}p$,

$\mathrm{E}[X]=1/p$, $\mathrm{Var}(X)=q/p^2$, $M_X(t)=pe^t/(1-qe^t)$.

NegBinom: $\mathrm{E}[Y]=r/p$, $\mathrm{Var}(Y)=qr/p^2$ (rth success)

$f(y) = \binom{y-1}{r-1}q^{y-r}p^r$,

Time Until rth success. FFFSSFS = (3,p)=7

Poisson: $\mathrm{E}[X]=\lambda=\mathrm{Var}(X)$, $M_X(t)=e^{\lambda(e^t-1)}$.

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

Lambda = arrivals in t time period. E.g. 4/hour

$X \sim \mathrm{Uniform}(a,b)$. $f(x)=\frac{1}{b-a}$ for $a\le x\le b$, $\mathrm{E}[X]=\frac{a+b}{2}$, $\mathrm{Var}(X)=\frac{(b-a)^2}{12}$, $M_X(t)=(e^{tb}-e^{ta})/(tb-ta)$.

$X \sim \mathrm{Exponential}(\lambda)$. $f(x)=\lambda e^{-\lambda x}$ for $x\ge 0$, $\mathrm{E}[X]=1/\lambda$, $\mathrm{Var}(X)=1/\lambda^2$, $M_X(t)=\lambda/(\lambda-t)$ for $t<\lambda$.

Theorem: The Exp(λ) has the *memoryless property*, i.e., for $s,t>0$, $P(X>s+t|X>s)=P(X>t)$.

Gamma: $f(x) = \frac{\lambda^\alpha x^{\alpha-1}e^{-\lambda x}}{\Gamma(\alpha)}$, $F_Y(y) = 1-e^{-\lambda y}\sum_{j=0}^{n-1}\frac{(\lambda y)^j}{j!}$, $y\ge 0$.

$\mathrm{E}[X]=\alpha/\lambda$, $\mathrm{Var}(X)=\alpha/\lambda^2$, $M_X(t)=\left[\lambda/(\lambda-t)\right]^\alpha$ for $t<\lambda$.

Gammas = Erlang = sum of K exponential variables.

$X \sim \mathrm{Beta}(a,b)$. $f(x)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$, $0\le x\le 1$, $a,b>0$.

$$\mathrm{E}[X] = \frac{a}{a+b} \quad\text{and}\quad \mathrm{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

$X \sim \mathrm{Normal}(\mu,\sigma^2)$. Most important distribution.

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

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

Theorem: If $X\sim\mathrm{Nor}(\mu,\sigma^2)$, then $aX+b\sim\mathrm{Nor}(a\mu+b,a^2\sigma^2)$.

Theorem: If X_1 and X_2 are *independent* with $X_i\sim\mathrm{Nor}(\mu_i,\sigma_i^2)$, $i=1,2$, then $X_1+X_2\sim\mathrm{Nor}(\mu_1+\mu_2,\ \sigma_1^2+\sigma_2^2)$.

Central Limit Theorem: If $X_1,X_2,\ldots,X_n \overset{\text{iid}}{\sim} f(x)$ with mean μ and variance σ^2, then

$$Z_n \equiv \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} = \frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} \xrightarrow{d} \mathrm{Nor}(0,1).$$

Statistics:

$$\mathrm{E}[\bar{X}] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n \mathrm{E}[X_i] = \mathrm{E}[X_i] = \mu.$$

So $\bar{X}$ is always unbiased for μ. That's why $\bar{X}$ is the *sample mean*.

Note: 1/X_bar is BIASES for lambda in the exponential case

$$\mathrm{E}[S^2] = \mathrm{E}\left[\frac{\sum_{i=1}^n(X_i-\bar{X})^2}{n-1}\right] = \mathrm{Var}(X_i) = \sigma^2.$$

Thus, S^2 is always unbiased for σ^2. This is why S^2 is called the *sample variance*.

$$\mathrm{MSE}(T) = \mathrm{Var}(T) + \underbrace{(\mathrm{E}[T]-\theta)}_{\text{Bias}}{}^2.$$

Lower MSA = Better, even if bias

Definition: The *bias* of an estimator $T(\mathbf{X})$ is $\mathrm{Bias}(T)\equiv\mathrm{E}[T]-\theta$.

The *mean squared error* of $T(\mathbf{X})$ is $\mathrm{MSE}(T)\equiv\mathrm{E}[(T-\theta)^2]$.

MLE

Definition: The *maximum likelihood estimator* (MLE) of θ is the value of θ that maximizes $L(\theta)$. The MLE is a function of the X_i's and is a RV.

$$L(\lambda) = \prod_{i=1}^n f(x_i)$$

Usually the E[X]. MLE wears little hat

$$\bar{X}_n - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}} \le \mu \le \bar{X}_n + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}},$$

(known)

$$\bar{X}_n - t_{\alpha/2,n-1}\sqrt{\frac{S^2}{n}} \le \mu \le \bar{X}_n + t_{\alpha/2,n-1}\sqrt{\frac{S^2}{n}},$$

unknown

Stepping through Dif EQ: $f(x+h) \approx f(x)+hf'(x)$ **. Iterate through.**

Steps in Simulations: Problem Formulation, Objectives and Planning, Model Building, Data Collection, Coding, Model Verification, Model Validation, Experimental Design, Run Experiments, Output Analysis, Make Reports / Imp.

System: Collection of entities that interact to accomplish a goal.

Model: Abstract representation of a system, uaually containing math / logical relationships.

System State: Set of variables that contain enough information to describe the system. Snapshot of the system

Entities: Can be permanent or temporary, have various properties or attributes.

List / Queue: ordered list of associated entities (a linked list, line of people)

Event: Point in time where the system state changes (arrival, departure, breakdown). Technically means the TIME something happens, but also WHAT.

Activity: duration of time of SPECIFIED length (unconditional wait)

Conditional wait: duration of time of unspecified length. Customer Wait Time

Simulation Clock: Variable representing simulated time. Can advance in fixed increment or next event time. Fixed increment used in continuous models (differential equations, and where data only available at fixed times). Next event is where you move to next most imminent event. Use Future Event List.

FELs: only change at event times. Nothing happens between. Inserts, deletes, moves around, or does nothing at each event.

Languages maintain FELs within them. Arena uses Process Interaction (better). Event scheduling is the bad one.

Process Interaction: Create, Process, Dispose.

Sim Languages: more than 100. Huge price range. Consider cost, ease of learning, world view, and features. Can learn: here, textbook, conferences, and vendor short courses.

Uniform Random Number Generation

Lousy Generators: Random Devices (Hard to store / repeat), random number tables (cumbersome, too small), Mid square method (below, positive serial correlation. Degenerates when Xi = 0003 eg), Fibonacci and Additive Congruential Generators (also below, small # follow small #)

Set *seed* $X_0 = 6632$; then $6632^2 \to 43\mathbf{9834}24$;

So $X_1 = 9834$; then $9834^2 \to 96\mathbf{7075}56$;

So $X_2 = 7075$, etc,...

$$X_i = (X_{i-1}+X_{i-2})\bmod m, \quad i=1,2,\ldots,$$

Linear Congruential Generators (LCGs, most widely used)

$X_i = (aX_{i-1}+c) \bmod m$, where X_0 is the seed.

$R_i = X_i/m$, $i=1,2,\ldots$

Choose a, c, m in a way to get good statistical quality and long period or cycle length (time until it repeats). If c=0, it's called multiplicative generator.

Full cycle generator is one where cycle length m = mod m.

Desert island: X0 between 1 and 2^(31) – 1. Cycle length > 2 billion

$$X_i \;=\; 16807\, X_{i-1} \bmod (2^{31}-1).$$

What can go wrong:

1. Not full period, only produces even integers
2. Produces non-random output
3. If m is small, quick cycling. Anything less than 2 billion or so
4. Just because M is big, you can still have subtle problems (RANDU)

RANDU:
$$X_i \;=\; 65539\, X_{i-1} \bmod 2^{31},$$

Tausworthe (more popular with computer science folks)

Define a sequence of binary digit *B1* B2 ,,... , by

$$B_i \;=\; \left(\sum_{j=1}^{q} c_j B_{i-j}\right) \bmod 2.$$

vvhere *c1* == 0 or 1. Looks a bit like a generalization of LCGs.

$$B_i \;=\; (B_{i-r}+B_{i-q}) \bmod 2 \;=\; B_{i-r} \text{ XOR } B_{i-q} \quad (0<r<q).$$

$$B_i = 0, \text{ if } B_{i-r} = B_{i-q} \quad \text{or} \quad B_i = 1, \text{ if } B_{i-r} \neq B_{i-q}.$$

Example (Law 2015): $r=3, q=5; B_1=\cdots=B_5=1$. Obtain
$B_i \;=\; (B_{i-3}+B_{i-5}) \bmod 2 \;=\; B_{i-3} \text{ XOR } B_{i-5}, \quad i>5$
$B_6 \;=\; (B_3 \text{ XOR } B_1) = 0,\; B_7 \;=\; (B_4 \text{ XOR } B_2) = 0$, etc.

1111 1000 1101 1101 0100 0010 0101 1001 1111 □

The period of 0-1 bits is always $2^q - 1 = 31$.

How do we go from B_i's to Unif(0,1)'s?

Easy way: Use (ℓ-bits in base 2)/2^ℓ and convert to base 10.

Example: Set $\ell = 4$ in previous example and get:

$$1111_2, 1000_2, 1101_2, 1101_2, \ldots \to \frac{15}{16}, \frac{8}{16}, \frac{13}{16}, \frac{13}{16}, \ldots . \quad \square$$

Bit conversion by base = base^I for I in range(len(#))

LCG Generalizations

A Simple Generalization:

$X_i = \left(\sum_{j=1}^{q} a_i X_{i-j}\right) \bmod m$, where the a_i's are constants.

Extremely large periods possible (up to $m^q - 1$ if parameters are chosen properly). But watch out! — Fibonacci is a special case.

Can combine two generators $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ to construct $Z_1, Z_2, \ldots$. Some suggestions:

- Set $Z_i = (X_i + Y_i) \bmod m$
- Shuffling
- Set $Z_i = X_i$ or $Z_i = Y_i$

L'Ecuyer generator has periods up to 2^191. Mersenne up to 2^(19937) -1.

You'll often need several billion PRNs, but never over 2^100

PRN Properties and Theory (meh)

Choosing a good generator – statistical tests

All tests have H0 as status quo, we reject with ample evidence against it.

Alpha (0.5 or 0.1) is the probably of a Type I error. P(Reject H0 | H0 is True)

Beta is probability of Type II error. P(Accept H0 | H0 is False)

Chi Squared GOF for Uniform numbers:

Test $H_0 : R_1, R_2, \ldots R_n \sim \text{Unif}(0,1)$.

Divide the unit interval into k cells (subintervals). If you choose equi-probable cells $[0, \frac{1}{k}), [\frac{1}{k}, \frac{2}{k}), \ldots, [\frac{k-1}{k}, 1]$, then a particular observation R_j will fall in a particular cell with prob $1/k$.

Tally how many of the n observations fall into the k cells. If $O_i \equiv$ # of R_j's in cell i, then (since the R_j's are i.i.d.), we can easily see that $O_i \sim \text{Bin}(n, \frac{1}{k})$, $i = 1, 2, \ldots, k$.

Thus, the expected number of R_j's to fall in cell i will be $E_i \equiv \mathrm{E}[O_i] = n/k$, $i = 1, 2, \ldots, k$.

$$\chi_0^2 \equiv \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.$$

Large Values indicate a bad fit.

In fact, we *reject* the null hypothesis H_0 (that the observations are uniform) if $\chi_0^2 > \chi^2_{\alpha,k-1}$, where $\chi^2_{\alpha,k-1}$ is the appropriate $(1-\alpha)$ quantile from a χ^2 table, i.e., $P(\chi^2_{k-1} < \chi^2_{\alpha,k-1}) = 1 - \alpha$.

If $\chi_0^2 \le \chi^2_{\alpha,k-1}$, we *fail to reject* H_0.

Usual recommendation from baby stats class: For the χ^2 g-o-f test to work, pick k, n such that $E_i \ge 5$ and n at least 30. But...

Unlike what you learned in baby stats class, when we test PRN generators, we usually have a *huge* number of observations n (at least millions) with a large number of cells k. When k is large, we can use the approximation

$$\chi^2_{\alpha,k-1} \;\approx\; (k-1)\left[1 - \frac{2}{9(k-1)} + z_\alpha \sqrt{\frac{2}{9(k-1)}}\right]^3,$$

where z_α is the appropriate standard normal quantile.

Independence Tests

Run: Series of similar observations. Run test rejects if there are too many or too few runs.

Up and Down test. If n is >= 20 and Rjs are independent, then:

$$A \;\approx\; \text{Nor}\left(\frac{2n-1}{3}, \frac{16n-29}{90}\right). \qquad Z_0 \;=\; \frac{A - \mathrm{E}[A]}{\sqrt{\mathrm{Var}(A)}},$$

Runs Test above and below the mean:

Fact: If n is large and the R_j's are actually independent, then

$$B \;\approx\; \text{Nor}\left(\frac{2n_1 n_2}{n} + \frac{1}{2}, \frac{2n_1 n_2(2n_1 n_2 - n)}{n^2(n-1)}\right),$$

where n_1 is the number of observations ≥ 0.5 and $n_2 = n - n_1$.

The test statistic is $Z_0 = (B - \mathrm{E}[B])/\sqrt{\mathrm{Var}(B)}$, and we reject H_0 if $|Z_0| > z_{\alpha/2}$.

Illustrative Example (from BCNN): Suppose that $n = 40$, with the following $+/-$ sequence.

$$- + + + + + + + - - - + + - + - - - - -$$
$$- - + + - - - - + + - - - + - + - - + + -$$

Then $n_1 = 18$, $n_2 = 22$, and $B = 17$. This implies that $\mathrm{E}[B] \doteq 20.3$ and $\mathrm{Var}(B) \doteq 9.54$. And this yields $Z_0 = -1.07$.

Since $|Z_0| < z_{\alpha/2} = 1.96$, we *fail* to reject the test; so we can treat the observations as independent. □

Random Variate GenerationI Using U(0,1) to generate variables)

Inverse Transform Theroem. If X is a continuous random variable, then F(X) ~ U(0,1).

- Sample U from U(0,1), then return X = F^{-1}(U)

Continuous Examples:

Weibull:
$$X = \frac{1}{\lambda}[-\ell n(1-U)]^{1/\beta} \quad \text{or} \quad X = \frac{1}{\lambda}[-\ell n(U)]^{1/\beta}.$$

Triangular:

If $U < 1/2$, we solve $X^2/2 = U$ to get $X = \sqrt{2U}$.
If $U \ge 1/2$, the only root of $1 - (X-2)^2/2 = U$ in $[1, 2]$ is
$$X = 2 - \sqrt{2(1-U)}.$$

$$Z \;=\; \Phi^{-1}(U) \;\approx\; \frac{U^{0.135} - (1-U)^{0.135}}{0.1975}.$$

Normal:

$$Z \;=\; \text{sign}(U - 1/2)\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right),$$

$$t \;=\; \{-\ell n[\min(U, 1-U)]^2\}^{1/2},$$

and

$$c_0 = 2.515517, \quad c_1 = 0.802853, \quad c_2 = 0.010328,$$
$$d_1 = 1.432788, \quad d_2 = 0.189269, \quad d_3 = 0.001308.$$

If you want X = Nor(mu, var), take X = mu + sqrt(var)*Z

$X \sim \text{Nor}(3, 16)$, and you start with $U = 0.59$. Then

$$X \;=\; \mu + \sigma Z \;=\; 3 + 4\Phi^{-1}(0.59) \;=\; 3 + 4(0.2275) \;=\; 3.91.$$

Discrete Examples (often best to construct a table):

$$X \;=\; \min[k : 1 - q^k \ge U] \;=\; \left\lceil \frac{\ell n(1-U)}{\ell n(1-p)} \right\rceil \;\sim\; \left\lceil \frac{\ell n(U)}{\ell n(1-p)} \right\rceil.$$

For instance, if $p = 0.3$ and $U = 0.72$, we obtain

$$X \;=\; \left\lceil \frac{\ell n(0.28)}{\ell n(0.7)} \right\rceil \;=\; 4. \quad \square$$

Geometric:

Can also generate a Geom by counting Bern(p) until success. Coin tosses, or # of generated Ui until Ui < p.

Remark: If you have a discrete distribution like Pois(λ) with an infinite number of values, you could write out table entries until the c.d.f. is nearly one, generate exactly one U, and then search until you find $X = F^{-1}(U)$, i.e., x_i such that $U \in (F(x_{i-1}), F(x_i)]$.

Convolution Method: Summing up things (bern -> binomial).

Binomical (3, 0.4). Ui = .63, U2=.17, u3 - .81, then Y=0+1+0 =1

Triangular: if U1 and U2 are iid U(0,1) then U1 + U2 is Tria(0,1,2)

Erlang:
$$Y \;=\; \sum_{i=1}^{n} X_i \;=\; \sum_{i=1}^{n}\left[\frac{-1}{\lambda}\ell n(U_i)\right] \;=\; \frac{-1}{\lambda}\ell n\left(\prod_{i=1}^{n} U_i\right).$$

CRUDE desert island Nor(0,1) generator:

Suppose that $U_1, \ldots, U_n$ are i.i.d. $\mathcal{U}(0,1)$, and let $Y = \sum_{i=1}^{n} U_i$.
For large n, the CLT implies that $Y \approx \text{Nor}(n/2, n/12)$.

In particular, let's choose $n = 12$, and assume that it's "large." Then

$$Y - 6 \;=\; \sum_{i=1}^{12} U_i - 6 \;\approx\; \text{Nor}(0, 1). \quad \square$$

If $X_1, \ldots, X_n$ are i.i.d. Geom(p), then $\sum_{i=1}^n X_i \sim \text{NegBin}(n, p)$.

If $Z_1, \ldots, Z_n$ are i.i.d. Nor(0,1), then $\sum_{i=1}^n Z_i^2 \sim \chi^2(n)$.

If $X_1, \ldots, X_n$ are i.i.d. Cauchy, then $\bar{X} \sim$ Cauchy (this is kind of like getting nowhere fast!).

Acceptance Rejection Method. Samples from a distribution that is "almost" the one we want, and adjusts by only acception a certain proportion.
For Uniform(2/3, 1), generate U(0,1) and accept is > 2/3

Notation: Suppose we want to simulate a continuous RV X with p.d.f. $f(x)$, but that it's difficult to generate directly. Also suppose that we can easily generate a RV having p.d.f. $h(x) \equiv t(x)/c$, where $t(x)$ *majorizes* $f(x)$, i.e.,

$$t(x) \geq f(x), \quad x \in \mathbb{R},$$

and

$$c \equiv \int_{-\infty}^{\infty} t(x)\, dx \geq \int_{-\infty}^{\infty} f(x)\, dx = 1,$$

where we assume that $c < \infty$.

Theorem (von Neumann 1951): Define $g(x) \equiv f(x)/t(x)$ and note that $0 \leq g(x) \leq 1$ for all x. Let $U \sim \mathcal{U}(0,1)$, and let Y be a RV (independent of U) with p.d.f. $h(y) = t(y)/c$. If $U \leq g(Y)$, then Y has (conditional) p.d.f. $f(y)$.

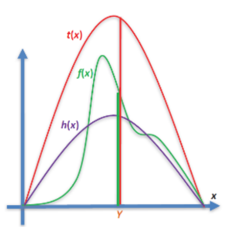
Repeat

Generate U from $\mathcal{U}(0,1)$

Generate Y from $h(y)$ (independent of U)

until $U \leq g(Y) = \dfrac{f(Y)}{t(Y)} = \dfrac{f(Y)}{c\,h(Y)}$

Return $X \leftarrow Y$



Generate a point Y uniformly under $t(x)$ (equivalently, sample Y from p.d.f. $h(x)$).

Accept the point with probability $f(Y) / t(Y) = f(Y) / [c\, h(Y)]$.

If you accept, then set $X = Y$ and stop.

Two key items: ability to sample quickly from h(y), and c must be small (t(x) is close to f(x). The # of trialcs until success is Geom(1/c), so mean # of trials is c

Example (Ross): Generate a standard *half-normal* RV, with p.d.f.

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \geq 0.$$

Use the majorizing function

$$t(x) = \sqrt{\frac{2e}{\pi}} e^{-x} \geq f(x), \quad \text{with} \quad c = \int_0^{\infty} t(x)\, dx = \sqrt{\frac{2e}{\pi}}.$$

Then

$$h(x) = t(x)/c = e^{-x} \quad \text{(easy Exp(1) p.d.f.)},$$

and

$$g(x) = f(x)/t(x) = e^{-(x-1)^2/2}. \quad \square$$

We can use the half-normal result to generate a Nor(0,1) variate.

Generate U from $\mathcal{U}(0,1)$.

Generate X from the half-normal distribution.

Return

$$Z = \begin{cases} -X & \text{if } U \leq 1/2 \\ X & \text{if } U > 1/2. \end{cases}$$

Poisson:

Algorithm

$a \leftarrow e^{-\lambda}$; $p \leftarrow 1$; $X \leftarrow -1$

Until $p < a$

Generate U from $\mathcal{U}(0,1)$

$p \leftarrow pU$; $X \leftarrow X + 1$

Return X

Sample until $e^{-\lambda} = 0.1353 > \prod_{i=1}^{n+1} U_i$.

n	U_{n+1}	$\prod_{i=1}^{n+1} U_i$	Stop?
0	0.3911	0.3911	No
1	0.9451	0.3696	No
2	0.5033	0.1860	No
3	0.7003	0.1303	Yes

Thus, we take $X = 3$. $\square$

Remark: An easy argument says that the expected number of U's that are required to generate one realization of X is $\mathrm{E}[X+1] = \lambda + 1$.

Remark: If $\lambda \geq 20$, we can use the normal approximation

$$\frac{X - \lambda}{\sqrt{\lambda}} \approx \text{Nor}(0,1).$$

Algorithm (for $\lambda \geq 20$)

Generate Z from Nor(0, 1).

Return $X = \max(0, \lfloor \lambda + \sqrt{\lambda} Z + 0.5 \rfloor)$ ("continuity correction").

E.g., if $\lambda = 30$ and $Z = 1.46$, then $X = \lfloor 30.5 + \sqrt{30}(1.46) \rfloor = 38$.

Composition:

The goal is to generate a RV with c.d.f.

$$F(x) = \sum_{j=1}^{\infty} p_j F_j(x),$$

- Generate a positive integer J such that $P(J = j) = p_j$ for all j.
- Return X from c.d.f. $F_J(x)$.

Theorem: If U_1, U_2 are i.i.d. $\mathcal{U}(0,1)$, then

$$\begin{aligned} Z_1 &= \sqrt{-2\ell n(U_1)} \cos(2\pi U_2) \\ Z_2 &= \sqrt{-2\ell n(U_1)} \sin(2\pi U_2) \end{aligned}$$

Box Muller: are i.i.d. Nor(0,1).

$$\frac{Z}{\sqrt{Y/n}} \sim t(n).$$

Example: Note that $Z_1^2 + Z_2^2 \sim \chi^2(1) + \chi^2(1) \sim \chi^2(2)$.

$\chi^2(2) \sim \text{Exp}(1/2)$.

Similarly, $Z_2^2/Z_1^2 = \tan^2(2\pi U) \sim t^2(1) \sim F(1,1)$.

Example: $Z_2/Z_1 \sim \text{Nor}(0,1)/\text{Nor}(0,1) \sim \text{Cauchy} \sim t(1)$. $\tan(2\pi U) \sim \text{Cauchy}$ (and, similarly, $\cot(2\pi U) \sim \text{Cauchy}$).

Order Statistics:

CDF Min Stat: $1 - [1 - F(y)]^n$. **CDF Max: [F(Y)]^n**

Thus, $Y = \min_i\{X_i\} \sim \text{Exp}(n\lambda)$. So take $Y = -\frac{1}{n\lambda} \ell n(U)$.

Max PRN for Exp(lamb) = -1/lamb * ln(1-U^(1/n))

MORE ARENA: You can get clocks, calendars, histograms, graphs

In Arena variables ARE global.

Record module: collects information about an entity when it passes through.

Batch: can make permanent or temporary. Need to split temp.

Split: duplicate for permanent batches, or use split existing for temp.

CAN assign service times as an attribute. Can also record arrival times thisway

CAN use logical expression for service time (if entity=A)*TRIA(1,3,4) etc.

CAN use fake customers to accomplish tasks (calculate probs, track time, breakdown demon).

Advanced Process Template: Delay, Seize, Release, Expression, Failure

Separate Delay, Seize, Release modules for more advanced actions.

CAN create failures with breakdown demons or Failure spreadsheet. Can do it by count or time. Failures can be Ignore (reduce repair time), Wait (delay repair) or preempt (complete current service after repair)

Blocks: Seize, Delay, Release, Queue, Alter. Certain primitive blocks can't connect to advanced or normal process module, so included in blocks.

Sets: Resource Sets, Counter, Tally, Entity, Entity Picture etc. Set spreadsheet in basic process. Resource set can have distinct servers with different schedules. Usually set preferred order of resource sets. CAREFUL: Make sure you release the same person you originally sieze'd. (Save in Index usually)

Set selection rules: Cyclical, Random, Preferred Order, Specific Member, largest remaining capacity, smallest number busy.

Call center examples: Nonhomogeneous Poisson Process.

Re-entrant Queues: going back to the same server. Can have different service times / priorities each time.

Advanced Transfer: Route, Station, Enter, Leave modules. Advanced Sets (Sets of sequences). Parts can move by themselves, via transporters, via conveyers (requires construction of transporter and conveyor paths)/

Station: Where you can go, **Route**: tells you where you're going, **Enter**: Way into a station, **Leave**: depart a station.

Sequence: Path through the system. Different types of customers / parts follow different sequences. (May have different servict time at each station depending on type and place in sequence). Sequence -> Steps -> Assignments (service distribution).

Advanced Sets: Sets of sequences.

Resource: Seize – Release, Transporter: Request-Free. Needs Distance Set.

Conveyor: Access-Exit. Requires Segment Set.

Multivariate Normal Generation

The random vector $\boldsymbol{X} = (X_1, \ldots, X_k)^{\mathrm{T}}$ has the *multivariate normal distribution* with mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)^{\mathrm{T}}$ and $k \times k$ covariance matrix $\Sigma = (\sigma_{ij})$ if it has p.d.f.

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left\{-\frac{(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{2}\right\}, \quad \boldsymbol{x} \in \mathbb{R}^k.$$

$$\mathrm{E}[X_i] = \mu_i, \quad \mathrm{Var}(X_i) = \sigma_{ii}, \quad \mathrm{Cov}(X_i, X_j) = \sigma_{ij}.$$

Notation: $\boldsymbol{X} \sim \mathrm{Nor}_k(\boldsymbol{\mu}, \Sigma)$.

In order to generate $\boldsymbol{X}$, let's start with a vector $\boldsymbol{Z} = (Z_1, \ldots, Z_k)$ of i.i.d. Nor(0,1) RV's. That is, suppose $\boldsymbol{Z} \sim \mathrm{Nor}_k(\boldsymbol{0}, I)$, where I is the $k \times k$ identity matrix, and $\boldsymbol{0}$ is simply a vector of 0's.

Suppose we can find the (lower triangular) Cholesky matrix C such that $\Sigma = CC^T$.

Then it can be shown that $\boldsymbol{X} = \boldsymbol{\mu} + C\boldsymbol{Z}$ is multivariate normal with mean $\boldsymbol{\mu}$ and covariance matrix

$$\Sigma \equiv \mathrm{E}[(C\boldsymbol{Z})(C\boldsymbol{Z})^T] = \mathrm{E}[C\boldsymbol{Z}\boldsymbol{Z}^T C^T] = C(\mathrm{E}[\boldsymbol{Z}\boldsymbol{Z}^T])C^T = CC^T$$

For $k = 2$, we can easily derive

$$C = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}} & \sqrt{\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}} \end{pmatrix}.$$

Since $\boldsymbol{X} = \boldsymbol{\mu} + C\boldsymbol{Z}$, we have

$$X_1 = \mu_1 + \sqrt{\sigma_{11}}\, Z_1$$
$$X_2 = \mu_2 + \frac{\sigma_{12}}{\sqrt{\sigma_{11}}} Z_1 + \sqrt{\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}}\, Z_2$$

Stochastic Processes

Markov Chains – use probability transitionmatrix. Can also use uniform

Poisson arrivals

When the arrival rate is a *constant* λ, the interarrivals of a Poisson(λ) process are i.i.d. Exp(λ), and the arrival times are:

$$T_0 \leftarrow 0 \quad \text{and} \quad T_i \leftarrow T_{i-1} - \frac{1}{\lambda}\ell n(U_i), \quad i \geq 1.$$

Now suppose that we want to generate a *fixed number* n of PP(λ) arrivals in a *fixed time interval* $[a, b]$. To do so, we note a theorem stating that the joint distribution of the n arrivals is the same as the joint distribution of the order statistics of n i.i.d. $\mathcal{U}(a, b)$ RV's.

Generate i.i.d. $U_1, \ldots, U_n$ from $\mathcal{U}(0, 1)$

Sort the U_i's: $U_{(1)} < U_{(2)} < \cdots < U_{(n)}$

Set the arrival times to $T_i \leftarrow a + (b - a)U_{(i)}$

Nonhomogeneous Poisson Processes

Example: Suppose that the arrival pattern to the Waffle House over a certain time period is a NHPP with $\lambda(t) = t^2$. Find the probability that there will be exactly 4 arrivals between times $t = 1$ and 2.

First of all, the number of arrivals in that time interval is

$$N(2) - N(1) \sim \mathrm{Pois}\left(\int_1^2 t^2\, dt\right) \sim \mathrm{Pois}(7/3).$$

Thus,

$$P(N(2) - N(1) = 4) = \frac{e^{-7/3}(7/3)^4}{4!} = 0.120. \quad \square$$

Use Thinning Algorithm. Assumes maxt lamba is finitie and does accept reject

Time Series Processes

MA(1): $$Y_i = \varepsilon_i + \theta\varepsilon_{i-1}, \quad \text{for } i = 1, 2, \ldots,$$

The MA(1) has covariance function $\mathrm{Var}(Y_i) = 1 + \theta^2$,

An AR(1) process is defined by

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad \text{for } i = 1, 2, \ldots,$$

where $-1 < \phi < 1$, $Y_0 \sim \mathrm{Nor}(0, 1)$, and the ε_i's are i.i.d. $\mathrm{Nor}(0, 1 - \phi^2)$ RV's that are independent of Y_0.

The AR(1) has covariance function $\mathrm{Cov}(Y_i, Y_{i+k}) = \phi^{|k|}$ for all $k = 0, \pm 1, \pm 2, \ldots$.

Multiply Nor(0,1) times sqrt(1-phi^2) to get the e terms.

$$Y_i = \sum_{j=1}^{p} \phi_j Y_{i-j} + \epsilon_i + \sum_{j=1}^{q} \theta_j \epsilon_{i-j}, \quad i = 1, 2, \ldots,$$

ARMA combines AR and MA

An EAR(1) process (Lewis 1980) is defined by

$$Y_i = \begin{cases} \phi Y_{i-1}, & \text{w.p. } \phi \\ \phi Y_{i-1} + \varepsilon_i, & \text{w.p. } 1 - \phi \end{cases},$$

for $i = 1, 2, \ldots$, where $0 \leq \phi < 1$, $Y_0 \sim \mathrm{Exp}(1)$, and the ε_i's are i.i.d. Exp(1) RV's that are independent of Y_0.

Now feed the correlated U_i's into the inverse of the Pareto c.d.f. to obtain correlated Pareto RV's:

$$X_i = F_X^{-1}(U_i) = F_X^{-1}(\Phi(Y_i)) = \frac{\lambda}{[1 - \Phi(Y_i)]^{1/\beta}}, \quad i = 1, 2, \ldots.$$

Pareto:

$$W_{i+1}^Q = \max\{W_i^Q + S_i - I_{i+1},\ 0\}.$$

And similarly, the total time in system, $W_i = W_i^Q + S_i$, is

$$W_{i+1} = \max\{W_i - I_{i+1},\ 0\} + S_{i+1}.$$

Queuing

Brownian Motion: W(b) – W(a). W(0) = 0, W(t) is Nor(0,t). CLT is a special case of Brownian motion. Yi are iid with mean 0 variance 1

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} Y_i \xrightarrow{d} \mathcal{W}(t) \quad \text{as } n \to \infty,$$

where $\xrightarrow{d}$ denotes convergence in distribution as n gets big, and $\lfloor \cdot \rfloor$ is the floor function, e.g., $\lfloor 3.7 \rfloor = 3$.

- BM is continuous everywhere, but has no derivatives! (Deep!)
- $\mathrm{Cov}(\mathcal{W}(s), \mathcal{W}(t)) = \min(s, t)$.
- Area under $\mathcal{W}(t)$ is normal: $\int_0^1 \mathcal{W}(t)\, dt \sim \mathrm{Nor}(0, \frac{1}{3})$.
- A *Brownian bridge*, $\mathcal{B}(t)$, is conditioned BM such that $\mathcal{W}(0) = \mathcal{W}(1) = 0$.
- $\mathrm{Cov}(\mathcal{B}(s), \mathcal{B}(t)) = \min(s, t) - st$.
- $\int_0^1 \mathcal{B}(t)\, dt \sim \mathrm{Nor}(0, \frac{1}{12})$.

Input Analysis

Discrete: Bernoulli, Binom, Geom. Neg Binom, Poisson, Empirical.

Continuous: Uniform, Triang, Expo, Normal, Beta, Gamma, Weibull, Empirical

Point Estimation: Statistics (mean, Variance) are RV. Different values for different samples. Used to estimate some unknown parameters.

X bar and S² usual estimates for mean and variance. Expected value should equal parameter and have low variance.

X bar is always unbiased for u, which is why it's the sample mean.

Same for S2. NOTE S is biased for standard deviation sigma

But be careful.... $1/\bar{X}$ is *biased* for λ in this exponential case, i.e., $\mathrm{E}[1/\bar{X}] \neq 1/\mathrm{E}[\bar{X}] = \lambda$.

$$\mathrm{MSE}(T) = \mathrm{Var}(T) + (\underbrace{\mathrm{E}[T] - \theta}_{\text{Bias}})^2.$$

Definition: The *relative efficiency* of T_2 to T_1 is $\mathrm{MSE}(T_1)/\mathrm{MSE}(T_2)$. If this quantity is < 1, then we'd want T_1.

Definition: Consider an iid random sample $X_1, \ldots, X_n$, where each X_i has pdf/pmf $f(x)$. Further, suppose that θ is some unknown parameter from X_i. The *likelihood function* is $L(\theta) \equiv \prod_{i=1}^n f(x_i)$.

Definition: The *maximum likelihood estimator* (MLE) of θ is the value of θ that maximizes $L(\theta)$. The MLE is a function of the X_i's and is a RV.

MLE for mean is usually the sample mean. MLE for variance is the sample variance but with only n for the denominator.

Remark: Notice how close $\widehat{\sigma^2}$ is to the (unbiased) sample variance

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{n}{n-1}\widehat{\sigma^2}$$

$\widehat{\sigma^2}$ is a little bit biased, but it has slightly less variance than S^2.

MLE for upper limit of Uniform:

Subject to this constraint, $L(\theta) = 1/\theta^n$ is maximized at the smallest possible θ value, namely, $\hat{\theta} = \max_i X_i$.

Theorem (Invariance Property): If $\hat{\theta}$ is the MLE of some parameter θ and $h(\cdot)$ is a one-to-one function, then $h(\hat{\theta})$ is the MLE of $h(\theta)$.

Recall: The kth *moment* of a random variable X is

$$\mathrm{E}[X^k] = \begin{cases} \sum_x x^k f(x) & \text{if } X \text{ is discrete} \\ \int_{\mathbb{R}} x^k f(x)\, dx & \text{if } X \text{ is cts} \end{cases}$$

Goodness of Fit (Ei > 5 and n >= 30)

1. Divide the domain of $f(x)$ into k sets, say, $A_1, A_2, \ldots, A_k$ (distinct points if X is discrete or intervals if X is continuous).
2. Tally the actual number of observations that fall in each set, say, O_i, $i = 1, 2, \ldots, k$. If $p_i \equiv P(X \in A_i)$, then $O_i \sim \mathrm{Bin}(n, p_i)$.
3. Determine the expected number of observations that would fall in each set if H_0 were true, say, $E_i = \mathrm{E}[O_i] = np_i$, $i = 1, 2, \ldots, k$.
4. Calculate a test statistic based on the differences between the E_i's and O_i's. The chi-squared g-o-f test statistic is
$$\chi_0^2 \equiv \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.$$
5. A large value of χ_0^2 indicates a bad fit.

We *reject* H_0 if $\chi_0^2 > \chi^2_{\alpha, k-1-s}$, where

- s is the number of unknown parameters from $f(x)$ that have to be estimated. E.g., if $X \sim \mathrm{Nor}(\mu, \sigma^2)$, then $s = 2$.
- $\chi^2_{\alpha,\nu}$ is the $(1 - \alpha)$ quantile of the χ^2_ν distribution, i.e., $P(\chi^2_\nu < \chi^2_{\alpha,\nu}) = 1 - \alpha$.

If $\chi_0^2 \leq \chi^2_{\alpha, k-1-s}$, we *fail to reject* H_0.

For continuous examples, you can divide into equal probability intervals.

Kolmogorov Smirnov GoF Test (More conservative)

The K-S test rejects H_0 if

$$D \equiv \max_{x \in \mathbb{R}} |F(x) - \hat{F}_n(x)| > D_{\alpha,n},$$

where α is the level of significance, and $D_{\alpha,n}$ is a K-S quantile that depends on the hypothesized c.d.f. $F(x)$.

X_i	0.039	0.706	0.016	0.198	0.793
$X_{(i)}$	0.016	0.039	0.198	0.706	0.793
$\frac{i}{n}$	0.2	0.4	0.6	0.8	1.0
$\frac{i-1}{n}$	0	0.2	0.4	0.6	0.8
$\frac{i}{n} - X_{(i)}$	0.184	0.361	**0.402**	0.094	0.207
$X_{(i)} - \frac{i-1}{n}$	0.016	–	–	**0.106**	–

Thus, $D^+ = 0.402$, $D^- = 0.106$, and then $D = 0.402$.
If we go to a K-S table for the uniform distribution, we have $D_{\alpha,n} = D_{0.05,5} = 0.565$. So we fail to reject uniformity. □

Input Analysis Issues: No/little data, Not a usual distribution, nonstationary data (distributions change), and multivariate/correlated data

Output Analysis: Interested in Mean, Variance, Quantiles, Probabilities etc. and want point estimators and Cis.

PROBLEM: Sims almost never produce raw output that is iid Normal.

Finite and Steady State Sims. When observations are not independent it impacts the Cis of the mean.

Since $\mathrm{E}[S_Y^2/n] \ll \mathrm{Var}(\bar{Y}_n)$, the CI will have true coverage $\ll 1-\alpha$! Oops! This is why you have to be really careful with correlated data!

SOLVE: Independent replications. Many independent runs of a simulation with their reported metrics used to estimate values. Avoids CI issues.

Denote the sample mean from replication i by

$$Z_i \equiv \frac{1}{m} \sum_{j=1}^{m} Y_{i,j},$$

where $Y_{i,j}$ is observation $j = 1, 2, \ldots, m$ from replication $i = 1, 2, \ldots, r$. E.g., $Y_{i,j}$ is customer j's waiting time from rep i.

Define the grand sample mean as $\bar{Z}_r \equiv \frac{1}{r}\sum_{i=1}^{r} Z_i$. The obvious point estimator for $\mathrm{Var}(\bar{Y}_m) = \mathrm{Var}(Z_i)$ is the sample variance of the Z_i's,

$$S_Z^2 \equiv \frac{1}{r-1} \sum_{i=1}^{r} (Z_i - \bar{Z}_r)^2.$$

Note that the forms of S_Z^2 and S_Y^2/m resemble each other. But since the replicate sample means are i.i.d., S_Z^2 is usually much less biased for $\mathrm{Var}(\bar{Y}_m) = \mathrm{Var}(Z_i)$ than is S_Y^2/m.

In light of the above, S_Z^2/r is a reasonable estimator for $\mathrm{Var}(\bar{Z}_r)$.

$$\theta \in \bar{Z}_r \pm t_{\alpha/2, r-1}\sqrt{S_Z^2/r}\,.$$

Where r = replications.

For smaller Cis: More replications. To reduce length of CI by factor of 10, need to increase reps by factor of 100. **(H/epsilon)^2*r**

Example: Let's suppose that we want a 95% CI for $\xi_{0.9}$ and we've made $r = 1000$ reps.

Point estimator for $\xi_{0.9}$ is $\widehat{\xi}_{0.9} = W_{(\lfloor 1000(0.9)+0.5 \rfloor)} = W_{(900)}$.

With the CI in mind,

$$j = \left\lfloor 900.5 - 1.96\sqrt{90} \right\rfloor = 881, \quad k = \left\lceil 900.5 + 1.96\sqrt{90} \right\rceil = 920,$$

so that the CI 95% is $[W_{(881)}, W_{(920)}]$. □

Initialization issues: test for bias in initialization visually, truncate after warmup time. Make a very long run to overwhelm initialization effects.

Batch means is often used in steady state analysis to calculate CI for mu. Divide long run into contiguous batches and apply CLT. Batch means are correlated for small m, but ok for large m.

Nevertheless, the story is that the batch size *m* needed to minimize MSE is O($n^{1/3}$). CI formula is the same as IR, but is one long run divided. Recommended to do 30 batches and increase batch size m.

Overlapping Batch Means has the same bias, but lower variance than regular batch means! Preferable if you have large m.

Use Regeneration: batch by when customer wait times hit zero.

Standardized Time series.

Confidence Intervals

One Sample:

$$\mu \in \bar{X}_n \pm t_{\alpha/2,n-1} S_X/\sqrt{n}.$$

Pooled CI: If the X's and Y's are independent but with *common, unknown variance*, then the usual CI for the difference in means is

$$\mu_X - \mu_Y \in \bar{X}_n - \bar{Y}_m \pm t_{\alpha/2,n+m-2}\, S_P \sqrt{\frac{1}{n} + \frac{1}{m}},$$

where

$$S_P^2 \equiv \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$$

is the pooled variance estimator for σ^2.

Approximate CI: If the X's and Y's are independent but with arbitrary unknown variances, then the usual CI for the difference in means is

$$\mu_X - \mu_Y \in \bar{X} - \bar{Y} \pm t_{\alpha/2,\nu} \sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}.$$

$$\nu \equiv \frac{\left(\frac{S_X^2}{n} + \frac{S_Y^2}{m}\right)^2}{\frac{(S_X^2/n)^2}{n+1} + \frac{(S_Y^2/m)^2}{m+1}} - 2.$$

Paired t-test: Where D=Differences in observations. Yields same mean value but reduces variance when Cov(X,Y) is fairly positive

$$\bar{D} \equiv \frac{1}{n}\sum_{i=1}^{n} D_i \sim \mathrm{Nor}(\mu_D, \sigma_D^2/n)$$

$$S_D^2 \equiv \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2 \sim \frac{\sigma_D^2 \chi^2(n-1)}{n-1}.$$

Just like before, get the CI

$$\mu_D \in \bar{D} \pm t_{\alpha/2,n-1}\sqrt{S_D^2/n}.$$

Can achieve pairing by using common random numbers. (sometimes means shift slightly compared to unpaired)

Antithetic random umbers

Antithetic Random Numbers

Opposite of CRN — Suppose that $\hat{\theta}_1$ and $\hat{\theta}_2$ are i.i.d. unbiased estimators for some parameter θ.

If we can induce *negative* correlation between $\hat{\theta}_1$ and $\hat{\theta}_2$, then the average of the two is also unbiased and may have very low variance,

Ranking and Selection Methods

Selects the best system or a subset including the best

Indifference-Zone Probability Requirement: For specified constants $(P^\star, \delta^\star)$ with $\delta^\star > 0$ and $1/k < P^\star < 1$, we require

$$\mathrm{P(CS)} \geq P^\star \quad \text{whenever} \quad \mu_{[k]} - \mu_{[k-1]} \geq \delta^\star. \tag{1}$$

The constant $\delta^\star$ can be thought of as the "smallest difference worth detecting."

Parameter configurations $\boldsymbol{\mu}$ satisfying $\mu_{[k]} - \mu_{[k-1]} \geq \delta^\star$ are in the *preference-zone* for a correct selection.

If $\mu_{[k]} - \mu_{[k-1]} < \delta^\star$, then you're in the *indifference-zone*.

Single Stage Procedure: get sample means from all populations, and select max. Key points come from selecting sample size from a table.

		$\delta^\star/\sigma$									
k	$P^\star$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
2	0.75	91	23	11	6	4	3	2	2	2	1
	0.90	329	83	37	21	14	10	7	6	5	4
	0.95	542	136	61	34	22	16	12	9	7	6
	0.99	1083	271	121	68	44	31	23	17	14	11
3	0.75	206	52	23	13	9	6	5	4	3	3
	0.90	498	125	56	32	20	14	11	8	7	5
	0.95	735	184	82	46	30	21	15	12	10	8
	0.99	1309	328	146	82	53	37	27	21	17	14
4	0.75	283	71	32	18	12	8	6	5	4	3
	0.90	602	151	67	38	25	17	13	10	8	7
	0.95	851	213	95	54	35	24	18	14	11	9
	0.99	1442	361	161	91	58	41	30	23	18	15

Limits: assumes known variances and common. Conservative for sample size. Requires independent populations. Rinnots and Nelson procedures estimate variance in stage 1 then move on.

Selecting Bernoulli with largest p Table

		$P^\star$						
k	$\Delta^\star$	0.60	0.75	0.80	0.85	0.90	0.95	0.99
3	0.10	20	52	69	91	125	184	327
	0.20	5	13	17	23	31	46	81
	0.30	3	6	8	10	14	20	35
	0.40	2	4	5	6	8	11	20
	0.50	2	3	3	4	5	7	12
4	0.10	34	71	90	114	150	212	360
	0.20	9	18	23	29	38	53	89
	0.30	4	8	10	13	17	23	39
	0.40	3	5	6	7	9	13	21
	0.50	2	3	4	5	6	8	13

Smallest n for $\mathcal{B}_{\mathrm{SH}}$ to Guarantee Probability Requirement

Curtailment: stop when second place can at best tie.

$P^\star$	$\theta^\star$	$k=2$	$k=3$	$k=4$	$k=5$
0.75	2.0	5	12	20	29
	1.8	5	17	29	41
	1.6	9	26	46	68
	1.4	17	52	92	137
	1.2	55	181	326	486
0.90	2.0	15	29	43	58
	1.8	19	40	61	83
	1.6	31	64	98	134
	1.4	59	126	196	271
	1.2	199	437	692	964
0.95	2.0	23	42	61	81
	1.8	33	59	87	115
	1.6	49	94	139	185
	1.4	97	186	278	374
	1.2	327	645	979	1331

Sample Size n for $\mathcal{M}_{\mathrm{BEM}}$ to Guarantee (3)