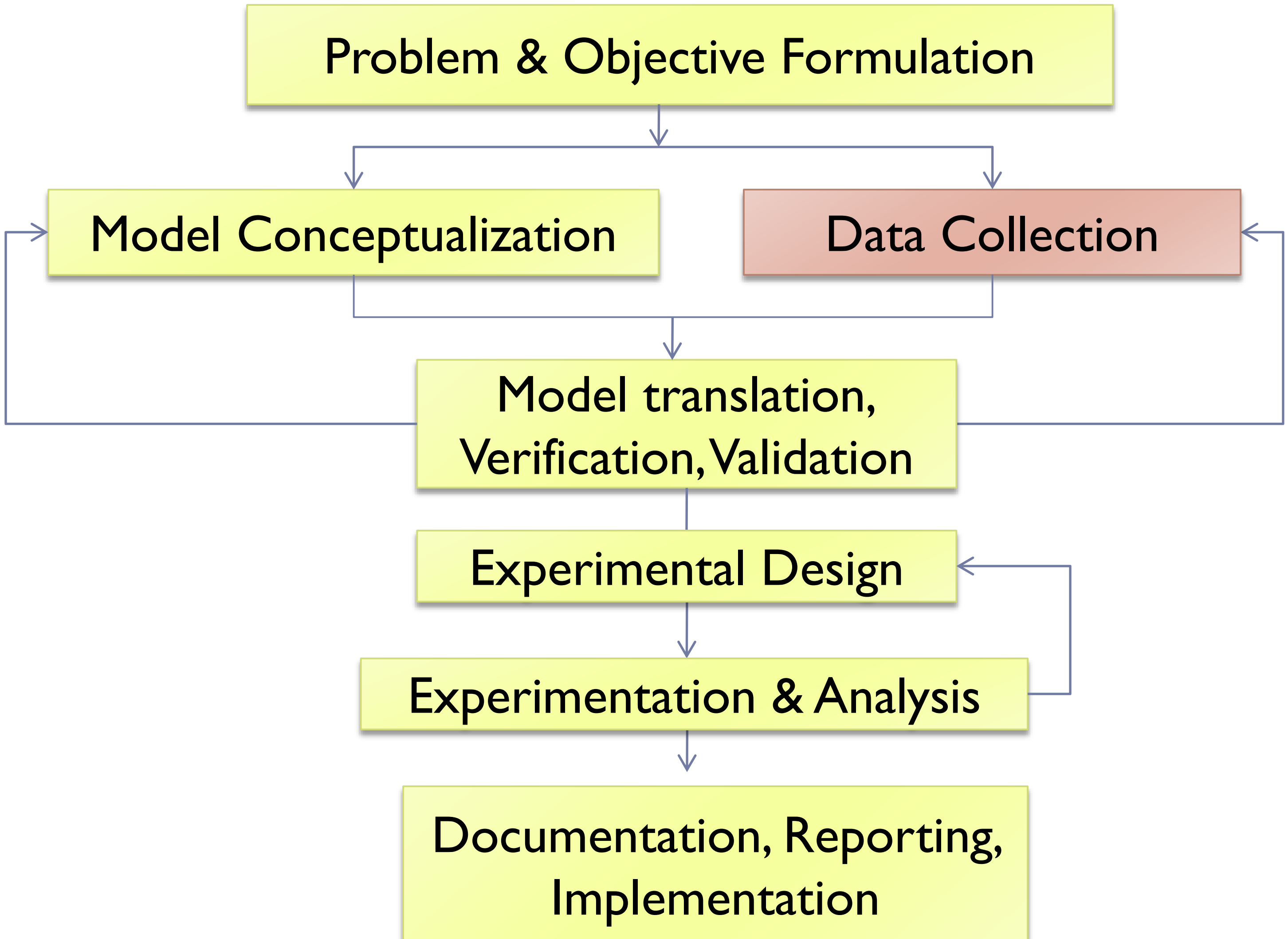
# Input Modeling

CS1538: Introduction to Simulations

# Steps in a Simulation Study



# Input Modeling

- In the real world, input data distributions are not always obvious or even clearly defined
  - We may not have any real data at all
    - If we are building a new network or road system, we wouldn't have a way to get the real data
  - We may only have a small number of sample data
    - If we can determine the distribution of the sample data, we might be able to generate enough for our simulation
  - The input data may not be from a single distribution
    - May differ at different times / days
  - Determining the distribution may not be easy
    - Usually requires multiple steps, and a combination of computer and "by hand" work

# No data at all: Create from scratch

- We would have to rely on knowledge about the problem
  - Experts
    - What do people in the area in question think, based on their knowledge and experience
  - Engineering specs
    - Ex: A device is built to have some mean time to failure based on the production environment. We can use that value as a starting point for mean time to failure of the device in the real environment
  - Similarity to something we already know
    - Ex: To figure out the input data distribution for a new road, we may be able to use data from similarly configured roads as a starting point

# Fitting sample data to a distribution

- Create one or more histograms of the data
- Graph them to see the basic "shape" of the distribution
  - How "wide" should each group be?
  - How do we know if a shape is of a particular distribution?
    - We may have some ideas about what might be a set of possible theoretical distributions
- Determine the parameters of the distribution
  - Ex: if we think it's exponential, we need to determine $\lambda$
  - Ex: if we think it's a normal, we need to determine $\mu$ and $\sigma$
- Apply goodness of fit tests
  - As with the random number testers, we can use the Chi-Square test and the Kolmogorov-Smirnov test

# Parameter Estimation

- Sample mean, X'
  - Average of the observed values (just like expectation)
- Sample variance
  - $S^2 = \Sigma (X_i - X')^2/(n-1) = ((\Sigma X_i^2) - nX'^2) / (n-1)$
  - Why divide by n-1?
    - This is necessary to keep the estimate unbiased – equally likely to over-estimate as under-estimate

# Parameter Estimation

- Suppose we believe that our empirical distribution is from the exponential family
  - If so, we would need to estimate $\lambda$
  - We have no prior belief about what value $\lambda$ should be
  - Pick $\lambda$ so that the chance of getting the empirical data is maximized
  - This is the Maximum Likelihood Estimate

# Maximum Likelihood Estimate (MLE)

- Suppose we observe data points $x_1, \ldots, x_n$. We want to pick $\lambda$ so that the observed data is the most likely to happen
  - $\text{argmax } \Pr(\lambda \mid x_1,\ldots,x_n) = \text{argmax } \Pr(x_1,\ldots,x_n \mid \lambda) \Pr(\lambda) / \Pr(x_1,\ldots,x_n)$
  - We can ignore the denominator because they are all the same
    - $\text{argmax } \Pr(\lambda \mid x_1,\ldots,x_n) = \text{argmax } \Pr(x_1,\ldots,x_n \mid \lambda) \Pr(\lambda)$
  - If we have no preference for any value of $\lambda$, then we can assume that $\Pr(\lambda)$ is from the uniform distribution so we can ignore it too
    - $\text{argmax } \Pr(\lambda \mid x_1,\ldots,x_n) = \text{argmax } \Pr(x_1,\ldots,x_n \mid \lambda)$
- True for any distribution (as long as assumptions are true)

# MLE for Exponentials

- For exponentials, we further know that
  - $Pr(x_1,\dots,x_n \mid \lambda) = f(x_1,\dots,x_n) = f(x_1)*\dots*f(x_n) = \lambda^n e^{-\lambda\Sigma xi}$
  - To argmax $Pr(x_1,\dots,x_n \mid \lambda)$, take the derivative; set to zero and solve for $\lambda$
    - First convert to log space $\ln f(\mathbf{x}) = n \ln \lambda - \lambda\Sigma x_i$
    - Take the derivative w/r/t $\lambda$ and set to 0: $n/\lambda - \Sigma x_i = 0$
    - So our estimate for $\lambda = 1/(\Sigma x_i /n) = 1/X'$
  - This makes intuitive sense because the expectation of an exponential distribution is $1/\lambda$, so our empirical estimation of $\lambda$ also has an inverse relationship with the empirical mean

# More parameter estimations

- What if we believe that the distribution is …
  - Binomial
    - $p' = X'/n$
  - Poisson
    - $\alpha' = X'$
  - Normal
    - $\mu' = X', \sigma' = S$
  - Gamma:
    - k : see Table A.9 (from Banks et al), $\theta' = 1/X'$

# Goodness of Fit Tests

- Once we have chosen a distribution and estimated its parameters, we need to check how well the distribution fits with the observed data.

- Goodness-of-fit tests
  - Chi-Squared Test (good for large sample sizes)
  - Kolmogorov-Smirnov Test

- The same general idea as when we checked for uniformity of the outputs from a PRNG
  - Except, a uniform distribution does not have any estimated parameters

# Goodness of Fit Tests

- Suppose expected distribution X has k possible outcomes. We compare the frequencies against the expected frequencies:
  - $C = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$

where $O_i$ is the observed number of occurrences of value $x_i$, and
$E_i$ is the expected number of occurrences of value $x_i$

- NULL Hypothesis, $H_0$ : O matches distribution of X
- Hypothesis $H_1$: O does not match distribution of X
- If the value for C is "too large" compared to the critical value, we reject the null hypothesis
- Critical value is determined by:
  - Degrees of freedom = k-s-1 where s is the number of sampled parameters
    - k is number of outcomes (also called bins)
    - s is the number of estimated parameters (e.g. for Exponentials, s=1; for Normals, s=2)
  - Level of significance

# Chi-Square Test Requirements

- Total number of observed data points > 20
- The expected frequency of each outcome is not too sparse (commonly required to be >= 5)
  - If $E_i < 5$ for some outcome $x_i$, it can be merged with an adjacent outcome
- If the distribution is continuous, bin the outcomes into k intervals
  - Set up the intervals for equal probability
  - To determine k, use the following guideline:

| Sample size (n) | Number of intervals (k) |
| --- | --- |
| <= 20 | N/A |
| 50 | 5-10 |
| 100 | 10-20 |
| > 100 | Sqrt(n) to n/5 |

# Example

- Suppose we observed these 50 time intervals between customer arrivals
- How should we begin to fit the observed data to a known distribution?

| | | | | |
|---|---|---|---|---|
| 5.409 | 0.028 | 15.31 | 1.641 | 3.83 |
| 5.933 | 2.025 | 19.00 | 8.533 | 7.349 |
| 12.75 | 6.167 | 1.291 | 6.333 | 3.899 |
| 6.314 | 10.63 | 0.389 | 1.833 | 6.59 |
| 12.19 | 15.12 | 0.322 | 13.45 | 8.192 |
| 0.263 | 6.777 | 4.523 | 8.793 | 13.85 |
| 7.33 | 2.31 | 11.57 | 1.25 | 16.53 |
| 15.84 | 31.25 | 6.863 | 29.22 | 11.35 |
| 7.552 | 0.962 | 10.47 | 2.32 | 7.207 |
| 0.985 | 0.939 | 6.45 | 0.532 | 4.238 |

(read column-first)

# Example

- Possible distribution choices
  - Poisson distribution (modeling the number of people arriving over some period of time)
    - Need to convert the 50 data points of inter-arrival times into cumulative time
    - Need to decide on a duration to use as a time period so that the number of outcomes (k) is appropriate
  - Exponential distribution (modeling the amount of time between arrivals)
    - When doing chi-square test, we'll need to quantize the continuous outcomes into bins of intervals where each interval has the same amount of probability mass

# Example: As Poisson Distribution

| | | | | |
|---|---|---|---|---|
| 5.409 | 74.596 | 166.085 | 228.600 | 304.689 |
| 11.342 | 76.622 | 185.087 | 237.134 | 312.038 |
| 24.092 | 82.789 | 186.378 | 243.467 | 315.937 |
| 30.407 | 93.420 | 186.767 | 245.300 | 322.527 |
| 42.599 | 108.541 | 187.089 | 258.746 | 330.718 |
| 42.862 | 115.318 | 191.611 | 267.539 | 344.573 |
| 50.192 | 117.628 | 203.180 | 268.790 | 361.105 |
| 66.031 | 148.877 | 210.043 | 298.006 | 372.457 |
| 73.582 | 149.839 | 220.509 | 300.326 | 379.664 |
| 74.567 | 150.777 | 226.959 | 300.859 | 383.902 |

If we want to treat the arrivals as a Poisson distribution, how should we go about analyzing the data?

# Processing the Observed Data using Poisson Distribution

- We really should have taken the data for a fixed duration multiple times
  - e.g., we might show up every weekday at Panera's and observe the number of customers who arrive between 4:00 and 4:15 for two weeks
- Since we took one long stream of observations, we'd be assuming that the rate of arrival is static
  - We can construct # of arrivals in some fixed period:
    - e.g., count how many people arrived in each 10 minute intervals

# Arrivals in 10-minute intervals

| | | | | |
|---|---|---|---|---|
| 5.409 | 74.596 | 166.085 | 228.600 | 304.689 |
| 11.342 | 76.622 | 185.087 | 237.134 | 312.038 |
| 24.092 | 82.789 | 186.378 | 243.467 | 315.937 |
| 30.407 | 93.420 | 186.767 | 245.300 | 322.527 |
| 42.599 | 108.541 | 187.089 | 258.746 | 330.718 |
| 42.862 | 115.318 | 191.611 | 267.539 | 344.573 |
| 50.192 | 117.628 | 203.180 | 268.790 | 361.105 |
| 66.031 | 148.877 | 210.043 | 298.006 | 372.457 |
| 73.582 | 149.839 | 220.509 | 300.326 | 379.664 |
| 74.567 | 150.777 | 226.959 | 300.859 | 383 |

| Period | # of Arrivals | Period | # of Arrivals | Period | # of Arrivals | Period | # of Arrivals |
|---|---|---|---|---|---|---|---|
| 10 | 1 | 110 | 1 | 210 | 1 | 310 | 3 |
| 20 | 1 | 120 | 2 | 220 | 1 | 320 | 2 |
| 30 | 1 | 130 | 0 | 230 | 3 | 330 | 1 |
| 40 | 1 | 140 | 0 | 240 | 1 | 340 | 1 |
| 50 | 2 | 150 | 2 | 250 | 2 | 350 | 1 |
| 60 | 1 | 160 | 1 | 260 | 1 | 360 | 0 |
| 70 | 1 | 170 | 1 | 270 | 2 | 370 | 1 |
| 80 | 4 | 180 | 0 | 280 | 0 | 380 | 2 |
| 90 | 1 | 190 | 4 | 290 | 0 | 390 | 1 |
| 100 | 1 | 200 | 1 | 300 | 1 | | |

# Arrivals

| Period | # of Arrivals | Period | # of Arrivals | Period | # of Arrivals | Period | # of Arrivals |
|---|---|---|---|---|---|---|---|
| 10 | 1 | 110 | 1 | 210 | 1 | 310 | 3 |
| 20 | 1 | 120 | 2 | 220 | 1 | 320 | 2 |
| 30 | 1 | 130 | 0 | 230 | 3 | 330 | 1 |
| 40 | 1 | 140 | 0 | 240 | 1 | 340 | 1 |
| 50 | 2 | 150 | 2 | 250 | 2 | 350 | 1 |
| 60 | 1 | 160 | 1 | 260 | 1 | 360 | 0 |
| 70 | 1 | 170 | 1 | 270 | 2 | 370 | 1 |
| 80 | 4 | 180 | 0 | 280 | 0 | 380 | 2 |
| 90 | 1 | 190 | 4 | 290 | 0 | 390 | 1 |
| 100 | 1 | 200 | 1 | 300 | 1 | | |

| Arrivals per Period | Frequency |
|---|---|
| 0 | 6 |
| 1 | 22 |
| 2 | 7 |
| 3 | 2 |
| 4+ | 2 |

Sample mean = #arrivals/#intervals
= 50/39 = 1.28
Sample variance = 0.945

Estimated rate of arrival is
1.28 person per 10 minutes, or
0.128 person per minute

# Example: Chi-Square Test for Poisson

| arrivals | Observed freq | Expected freq |
|---|---|---|
| 0 | 6.000 | 10.821 |
| 1 | 22.000 | 13.873 |
| 2 | 7.000 | 8.893 |
| 3 | 2.000 | 3.801 |
| 4+ | 2.000 | 1.612 |

Some bins are too small…
Which and why?

# Example: Chi-Square Test for Poisson

| arrivals | Observed freq | Expected freq |
|---|---|---|
| 0 | 6.000 | 10.821 |
| 1 | 22.000 | 13.873 |
| 2 | 7.000 | 8.893 |
| 3 | 2.000 | 3.801 |
| 4+ | 2.000 | 1.612 |

| Bin together arrivals with low freq | Observed Freq After binning | expected freq After binning | $(O-E)^2/E$ |
|---|---|---|---|
| 0 | 6.000 | 10.821 | 2.148 |
| 1 | 22.000 | 13.873 | 4.760 |
| 2+ | 11.000 | 14.305 | 0.764 |
| | | C = | 7.672 |

For bins too small, merge with adjacent bins until large enough (what is "large enough")?

# Example: Chi-Square Test for Poisson

| arrivals | Observed freq | Expected freq | Bin together arrivals with low freq | Observed Freq After binning | expected freq After binning | $(O-E)^2/E$ |
|---|---|---|---|---|---|---|
| 0 | 6.000 | 10.821 | 0 | 6.000 | 10.821 | 2.148 |
| 1 | 22.000 | 13.873 | 1 | 22.000 | 13.873 | 4.760 |
| 2 | 7.000 | 8.893 | 2+ | 11.000 | 14.305 | 0.764 |
| 3 | 2.000 | 3.801 | | | | |
| 4+ | 2.000 | 1.612 | | | | |
| | | | | | C = | **7.672** |

We used k=3 bins and we have one estimated parameter ($\lambda$=0.128) so our degrees of freedom=3-1-1 = 1

We'd look up the $\chi^2$ table for the appropriate critical value to compare.
$\chi^2_{1,\,0.05} = 7.88 > 7.672 = C$ – accept $H_0$ – awfully close though…

# Example

- Possible distribution choices
  - Poisson distribution (modeling the number of people arriving over some period of time)
    - Need to convert the 50 data points of inter-arrival times into cumulative time
    - Need to decide on a duration to use as a time period so that the number of outcomes (k) is appropriate
  - Exponential distribution (modeling the amount of time between arrivals)
    - When doing chi-square test, we'll need to quantize the continuous outcomes into bins of intervals where each interval has the same amount of probability mass

# Processing the Observed Data as Exponential Distribution

- Sample mean = 7.678
  - Average of observations
- Sample variance = 46.738
  - Variance of observations
- Estimated $\lambda = 1/7.678 = 0.130$

- So in our goodness of fit, we want to see how closely the observed dataset might have come from
  - $f(x) = 0.13\ e^{-0.13x}$

- Apply the chi-square test to check the goodness of fit

# Example: Chi-Square Test for Exponential

- To use Chi-Square Test, we need to decide on a way to bin the observed data so that we have frequency counts
  - Rather than fixed value intervals, we can define the intervals so that they have equal probability mass
  - In general, we still want the number of data points in each interval to be greater than 5
  - So if we have n data points, we might pick interval probability mass to be at least 5/n
    - In our example, each interval would have to take up at least 10% probability mass

# Example: Chi-Square Test for Exponential

- Let's say that we decided to break up the exponential function into 5 intervals (20% prob mass each)
  - CDF of our exponential is $F(x) = 1 - e^{-0.13x}$
  - Solve for x when F(x)=0.2, 0.4, etc.

# Example: Chi-Squared for Exponential

| | | | | |
|---|---|---|---|---|
| 5.409 | 0.028 | 15.31 | 1.641 | 3.83 |
| 5.933 | 2.025 | 19.00 | 8.533 | 7.349 |
| 12.75 | 6.167 | 1.291 | 6.333 | 3.899 |
| 6.314 | 10.63 | 0.389 | 1.833 | 6.59 |
| 12.19 | 15.12 | 0.322 | 13.45 | 8.192 |
| 0.263 | 6.777 | 4.523 | 8.793 | 13.85 |
| 7.33 | 2.31 | 11.57 | 1.25 | 16.53 |
| 15.84 | 31.25 | 6.863 | 29.22 | 11.35 |
| 7.552 | 0.962 | 10.47 | 2.32 | 7.207 |
| 0.985 | 0.939 | 6.45 | 0.532 | 4.238 |

(read column-first)

# What Next for Simulations?

- Once we've settled on a distribution that describes out input, what do we do with it?
  - Why were we attempting to model out input?
  - How do we use it in our simulations?

# Input Modeling

- To run a simulation, we need to be able to generate realistic input data

- Challenges:
  - We may not have any real data at all
  - We may only have a small number of sample data
  - Determining the distribution may not be easy
  - The input data may not be from a single distribution
    - Multiple variables
    - The input data may not be independent over time

# Multivariate and Time-Series Input Model

- If the input random variables are not independent, we need to be able to account for the dependence
- Multivariate input models:
  - The input is described by a fixed, finite number of random variables
  - Ex: The number of pedestrians arriving at an intersection and the number of cars arriving at an intersection
- Time-series input models:
  - A sequence of related random variables
  - Can be conceptually infinite
  - Ex: The size of the audience for a stage play over consecutive evenings

# Example 9.21 from Banks et al.

- A supply-chain simulation includes the lead time and annual demand for industrial robots. An increase in demand results in an increase in lead time: The final assembly of the robots must be made according to the specifications of the purchaser.
  - Therefore, rather than treat lead time and demand as independent random variables, a multivariate input model should be used

# Example 9.22 from Banks et al.

- A simulation of the web-based trading site of a stock broker includes the time between arrivals of orders to buy and sell.
  - Might be tempted to model inter-arrivals time naïvely using Exponential distribution. However, …

# Example 9.22 from Banks et al.

- A simulation of the web-based trading site of a stock broker includes the time between arrivals of orders to buy and sell.
  - Might be tempted to model inter-arrivals time naïvely using Exponential distribution. However, …
- Investors tend to react to what other investors are doing, so these buy and sell orders arrive in bursts.
  - Therefore, rather than treat the time between arrivals as independent random variables, a time series model should be developed.

# Example 9.22 from Banks et al.

- A simulation of the web-based trading site of a stock broker includes the time between arrivals of orders to buy and sell.
  - Might be tempted to model inter-arrivals time naïvely using Exponential distribution. However, …
- Investors tend to react to what other investors are doing, so these buy and sell orders arrive in bursts.
  - Therefore, rather than treat the time between arrivals as independent random variables, a time series model should be developed.

How is this different from non-stationary Poisson Process?

# Recall Covariance and Correlation

- Let X be a random variable with mean $\mu_X$ and variance $\sigma^2_X$ and let Y be a random variable with mean $\mu_Y$ and variance $\sigma^2_Y$
- The covariance between X and Y is defined to be
  - $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X\mu_Y$
  - If X and Y are independent, Cov(X,Y) = 0
    - Cov(X,X) = Var(X)
- The correlation between X and Y is defined to be
  - $\rho = Corr(X,Y) = Cov(X,Y)/\sigma_X\sigma_Y$
    - Corr(X,Y) = 0: X and Y are independent
    - 1> Corr(X,Y) > 0: they are positively correlated
    - -1 < Corr(X,Y) < 0: they are negatively correlated

# Useful Case: Bivariate Normal Distribution

- If X and Y are both normally distributed, the dependence between them can be modeled by the bivariate normal distribution with parameters $\mu_X$, $\mu_Y$, $\sigma^2_X$, $\sigma^2_Y$, and $\rho$ = Corr(X,Y)
  - We can estimate $\mu_X$, $\mu_Y$, $\sigma^2_X$, $\sigma^2_Y$ empirically from the sample data (X',Y', $S^2_X$, $S^2_Y$)
  - To estimate $\rho$, we would first need to estimate the covariance:
    - $Cov'(X,Y) = 1/(n-1)\ \Sigma_{j=1..n} (X_j - X')(Y_j - Y')$
      $= 1/(n-1)\ (\Sigma_{j=1..n} X_j Y_j) - nX'Y'$
    - $\rho' = Cov'(X,Y)/(S_X\ S_y)$

# Example 9.23
# (continuation of 9.21: Lead time & demand)

- Let L represent average lead time (in months)
  - $L' = 6.14$
  - $S_L' = 1.02$
- Let D represent annual demand
  - $D' = 101.80$
  - $S_D' = 9.93$
- Are L and D normally distributed?
  - How do we check?

| Lead Time | Demand |
|---|---|
| 6.5 | 103 |
| 4.3 | 83 |
| 6.9 | 116 |
| 6.0 | 97 |
| 6.9 | 112 |
| 6.9 | 104 |
| 5.8 | 106 |
| 7.3 | 109 |
| 4.5 | 92 |
| 6.3 | 96 |

# Example 9.23 (continuation of 9.21: Lead time & demand)

- $Cov'(L, D) = 1/(n-1)\ (\Sigma_{j=1..n} L_j D_j) - nL'D'$

  $= 1/(10-1)\ (\Sigma_{j=1..10} L_j D_j) - 10*6.14*101.80$

  $= 8.66$

- $\rho' = Cov'(L,D)/(S_L\ S_D) = 8.66/(1.02*9.93) = 0.86$

- $\rho' > 0$
  - L and D are positively correlated
- $\rho'$ close to 1
  - Strongly dependent

# Bivariate Normal Distribution

- We can generate more data points that follow a bivariate normal distribution:
  - Generate two independent standard normal random variables, $Z_1$ and $Z_2$
    - How do we do that?
  - Let $X = \mu_X + \sigma_X Z_1$
  - Let $Y = \mu_Y + \sigma_Y(\rho Z_1 + \text{sqrt}(1-\rho^2)Z_2)$
- For the example:
  - $L^* = L' + S_L' Z_1$
  - $D^* = D' + S_D'(\rho Z_1 + \text{sqrt}(1-\rho^2)Z_2)$
- Also possible:
  - *k*-variate normal distribution
  - Transform bivariate normal distribution to non-normal bivariate distributions

# Time-Series Input Models

- Time series is a sequence of random variables $X_1, X_2, X_3$, … that are identically distributed but could be dependent
  - $Cov(X_t, X_{t+h})$: lag-h autocovariance
  - $Corr(X_t, X_{t+h})$: lag-h autocorrelation
  - This measures the dependence between random variables that are separated by h-1 others in the time series
- If the value of the autocovariance depends only on h and not on t, then the time series is covariance stationary
  - $\rho_h = Corr(X_t, X_{t+h}) = \rho^h$
  - That is, the lag-h autocorrelation decreases geometrically as the lag increases
  - If the observations are far apart, they are nearly independent

# Example 9.22 from Banks et al.

- A simulation of the web-based trading site of a stock broker includes the time between arrivals of orders to buy and sell. Investors tend to react to what other investors are doing, so these buy and sell orders arrive in bursts.
  - Therefore, rather than treat the time between arrivals as independent random variables, a time series model should be developed.

# Example 9.24
# (continuation of 9.22: Stock broker)

- Suppose we have the 20 time gaps between customer buy and sell orders (in seconds) on the right
  - $T' = 5.2$ s
  - $S_T'^2 = 26.7\ s^2$

| Time between orders (sec) | |
|---|---|
| 1.95 | 0.68 |
| 1.75 | 0.61 |
| 1.58 | 11.98 |
| 1.42 | 10.79 |
| 1.28 | 9.71 |
| 1.15 | 14.02 |
| 1.04 | 12.62 |
| 0.93 | 11.36 |
| 0.84 | 10.22 |
| 0.75 | 9.20 |

(read column-first)

# Generating Random Variates of Time-Series Models

- Exponential Autoregressive Order-1 Model
  - Also called EAR(1) Model
  - Use EAR(1) if autocorrelation > 0

- Generate $T_t$ according to:
  - $T_t = \begin{cases} \phi * X_{t-1} & with\ probability\ \phi \\ \phi * X_{t-1} + \epsilon_t & with\ probability\ 1-\phi \end{cases}$
  - t = 2, 3, …
  - $\varepsilon_2, \varepsilon_3, \ldots$ are independent and identically (exponentially) distributed with mean $1/\lambda$
  - $0 <= \phi < 1$

- Estimate parameters:
  - $\phi' = \rho'$
  - $\lambda' = 1/X'$

# Generating (stationary) EAR(1) Time series

1. Generate $X_1$ from exponential distribution with mean $1/\lambda$
2. Set t = 2
3. Generate U from U[0, 1)
4. If U <= $\phi$:
   - $X_t = \phi * X_{t-1}$
5. Else:
   - Generate $\varepsilon_t$ from exponential with mean $1/\lambda$
   - $X_t = \phi * X_{t-1} + \varepsilon_t$
6. t += 1
7. Go to Step 3