MET3220C & MET6480 Computational Statistics

Parametric Probability Distributions

Discrete Distributions

Key Point: ALWAYS LOOK AT THE DATA !!!!
DOES THE DATA REALY FIT THE DISTRIBUTION

Parametric vs. Empirical Distributions

- Empirical distributions are based on a match to the sample data.
  - They are not based on underlying physical knowledge
- Parametric distributions are mathematical models (idealizations). In some cases the idealizations can be expected to be of very high quality.
  - One key question in later lectures will be 'how well does the distribution describe the data?'

Advantages of Parametric Distributions

- Compactness: The probability distribution of a parametric distribution can be explained with a formula.
  - Empirical distributions could require a very complicated histogram or pdf to describe the data.
- Smoothing and interpolation: Smoothing is largely unnecessary, or can be determined through integration of a formula. Interpolation is unnecessary.
  - For Empirical distributions, smoothing can be very complicated – or very misleading. Interpolation can be a nightmare in data spare parts of the distribution.
- Extrapolation: If the parametric distribution is believed to be sufficiently accurate for conditions outside values in the sample data, then extrapolation is simply a matter of working with the equation describing the parametric distribution.

Discrete vs. Continuous Distributions

- Parametric distributions can be classified as either discrete or continuous.
- Continuous distributions apply to data that can have any value (including fractions) between specific limits.
  - Example: Gaussian distribution has limit of $\pm\infty$.
- Discrete distributions contain only specific values.
  - Example: a binomial distribution has integer values from 0 to N, where N is the number of samples.
    - E.g., five heads out of nine coin flips.

Binomial Distribution

- A binomial distribution describes the probability of occurrence of the number of times a outcome occurs, given a number of samples and the probability of that outcome for a single trial.
- For N trials, there will be N+1 possible outcomes, range from zero occurrences to N occurrences.
- Two key considerations for applicability of a binomial distribution are:
  - 1) the probability of occurrence does not change from event to event, and
  - 2) The outcome of each trial is independent from the other outcomes.
- Close approximations to these conditions are often acceptable.

Binomial Distribution Example

- Used to describe the outcome of a certain number of elementary events, for which the elementary event has only two possible outcomes.
  - Example: the daily maximum temperature > 60°F
- These outcomes must be mutually exclusive.
- The events must be independent.
- The probability of the outcome of a single event must not change.
- Given N trials (N is the number of elementary events), there are N+1 possible outcomes.
  - E.g., 0, 1, 2, … , N – 1, or N days when T_{max} > 60°F.
  - Note that for this example, the samples would have to be separated by quite a few days to truly be independent.

Binomial Distribution Formula

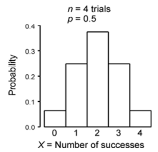
- Consider the probability of X of the N events occurring.
  - For example, X heads out of N coin flips.
- The probability of X events occurring (Pr{X=x}) is written as

$$\Pr\{X=x\} = \binom{N}{x} p^{x}(1-p)^{N-x}, \text{for } x = 1, 2, 3, \ldots, N$$

- Where p is the probability of occurrence, and (1 – p) is the probability of non-occurrence.

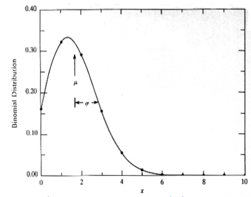
$$\binom{N}{x} = \frac{N!}{x!(N-x)!}$$

Where N ! = 1 × 2 × 3 × … × N

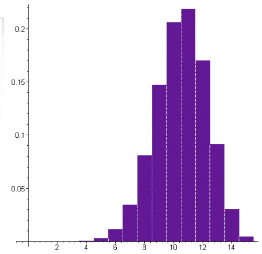


Graphic from www.zoology.ubc.ca/~bio300b/binomialnotes.html

More Examples



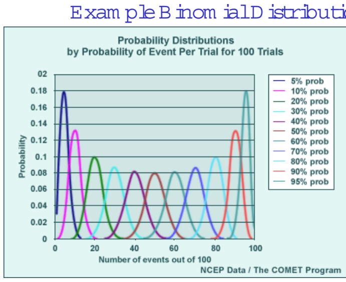
Graphic from www.physics.uwa.edu.au/.../node13.html



Graphic from http://adept.maplesoft.com/categories/education/statistics/html/images/Binomial_Distribution/Binomial_Distribution29.gif

- Consider the following categories for each of the above plots:
  - p < 0.5, p = 0.5, p > 0.5
  - Identify the category for each plot.

Example Binomial Distribution



- Note that the minimum in the peak occurs at p = 0.5.
- The spread of the distributions is largest for p = 0.5.

Graphic from http://meted.ucar.edu/nwp/pcu1/ensemble/media/graphics/binomial.gif

Binomial Distribution: Example

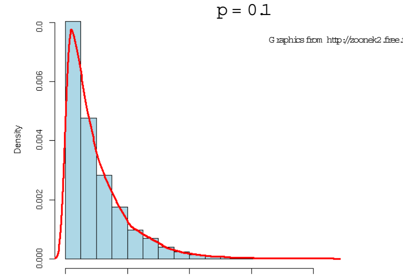
- The elementary event is the probability of a lake freezing in winter.
  - The probability in any year is 0.045
- Consider ten years of events.
- What is the probability of the lake freezing on at least one of those years?
  - Pr{X≥1} = Pr{X=1} + Pr{X=2} + Pr{X=3}+… + Pr{X=10}
  - Or in an easier form: Pr{X≥1} = 1 – Pr{X=0}
- Pr{X≥1} = 1 – Pr{X=0} = 1 – [10!/(0!10!)] $(0.045)^{0}$ $(0.955)^{10}$
  = 0.37

Geometric Distribution

- The geometric distribution is used to determine the probability of a successful outcome (in the last trial) in a fixed number of trials.
  - Otherwise the assumptions required to apply to the geometric distribution are identical to those required for a binomial distribution.
- The probability of a (one) successful trial in X events is
  - Pr{X=x} = p $(1 - p)^{x-1}$
- One application of this distribution is the length of waits or 'spells.' For example, wet spells or cold spells

Example Geometric Distribution

p = 0.1

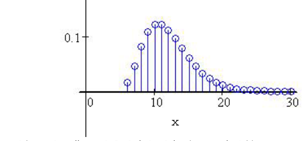


Graphics from http://zoonek2.free.fr/UNIX/48_R/g499.png

- Think about why there is a peak
  - Either probability of any success is small (left of peak), or
  - probability of a successes is greater than probability of no successes.

Negative Binomial Distribution

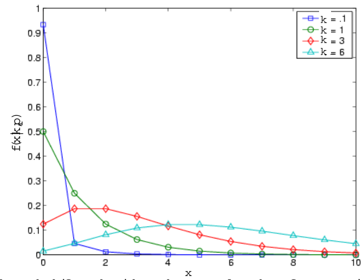
- The negative binomial distribution describes the probability of a number of failures (X), prior to the k^{th} success.
  - In other words, after X+k events, there were X failures in the first X+k-1 events. The last event must be the k^{th} success.
- $Pr\{X=x\} = \Gamma(k+x) / (x!\,\Gamma(k))\, p^k (1-p)^x$
  $= (k+x-1)! / (x!\,(k-1)!)\, p^k (1-p)^x$
- For typical applications k is an integer, in which case $\Gamma(k) = (k-1)!$



- Negative Binomial probability function with parameters k = 6, x = 24, p = 0.5

Graphic from http://home.ubalt.edu/ntsbarsh/Business-stat/negbi.jpg

Example Negative Binomial Distribution



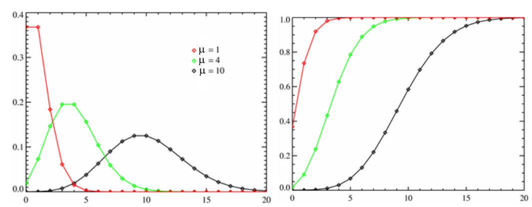
- The peak shifts to the right as the required number of successes increases. Why?
- Because more failures are likely to occur with the larger number of trials.

Graphic from http://www.mathworks.com/access/helpdesk/help/toolbox/stats/nbinpdf.gif

Poisson Distribution

- The Poisson distribution describes the likelihood of a certain number of events (x) occurring in a limited space and/or time.
  - As with the previous distributions, the events must be independent.
- It is the limiting case of the binomial distribution, when $p \to 0$ and $N \to \infty$
- Examples:
  - The number of hurricanes making landfall on the U.S. east coast in a year.
  - The annual number of tornadoes in Leon county
- The Poisson distribution is described by only one parameter: μ.
  - $Pr\{X=x\} = \mu^x e^{-\mu} / x!$, for x = 1, 2, 3, … .
  - This parameter (μ) is rather convenient because it is equal to both the standard deviation and the mean.

Example Poisson Distribution



- Poisson distributions are on the left, and the related cumulative probability distributions are on the right.
- Note that if you have a CDF, and a uniform random number generator, you can use the CDF to convert to the number associated with a randomly generated cumulative probability.

Graphic from http://en.wikipedia.org/wiki/Poisson_distribution

Other Characteristics of the Discrete Parametric Parameterizations

- Expectation value
  - Expectation value is a fancy way of saying the mean or average.
  - Square brackets are usually used to indicate an expectation value.
    - For example [x] (or E[x]) is the mean value of x.
- There are some simple rule that can be applied to products:
  - E[c] = c, where c is a constant
  - E[c g(x)] = c E[g(x)], where g(x) is a function of x.
  - $E[\Sigma g_j(x)] = \Sigma (E[g_j(x)])$

Distribution	$\mu = E(X)$	$\sigma^2 = Var[X]$
Binomial	N p	N p (1 – p)
Geometric	1/p	$(1-p)/p^2$
Negative Binomial	k (1 – p) /p	$K(1-p)/p^2$
Poisson	μ	μ