NREUP 2006: Probability and Networking

Dov Chelst

DeVry University

dchelst@nj.devry.edu

Five Discussions on Probability: Basic to Advanced

Teaching Outline (perhaps out of order!)

- Getting Started - Review & More!
  - Basic Probability
  - Discrete Distribution Formulas
  - Binomial, Poisson, etc.
  - Independence v. Dependence
  - Limits: Stirling's Formula
  - Counting: Binomial Coefficients
  - Taylor's Theorem & Taylor Series
  - "Big Oh" and "Little oh" Notation
- Modelling
  - Random Walks
  - Urn Models
  - Markov Chains:
    * Basic Description
    * Transience v. Recurrence
    * Absorption v. Stationarity
  - Matrices:
    * Multiplication
    * Eigenvalues & Eigenvectors
    * Applications

Definitions

- Sample Space $\mathcal{S}$: Collection of Outcomes
- Outcome x: Result of An Experiment
- Event A: Any portion of the sample space
- Random Variable X: A function that randomly produces one of out of a set of numbers following a well-defined rule.
- Probability P: A rule that assigns a numeric value (between 0 and 1) to every event in a sample space.
  - P is an area function in "space" whose total area is 1.
  - $0 \leq P(A) \leq 1$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - $P(A^c) = 1 - P(A)$

Discrete Distributions

- Discrete Random Variable X only takes on a finite (or countably infinite) set of values.
- This allows for a **sum** rather than an **integral**.
- Each outcome x is assigned a value $0 < p(x) < 1$ so that $\sum p(x) = 1$.
- Calculation Issues:
  - Determining the Exact Probabilities in Known Situations
  - Single (Density) v. Cumulative (Distribution) Probabilities
  - Difficult as N increases.
  - Determine Limits and Approximation Methods
- Application Issues:
  - When is a certain type distribution applicable?
  - What are each types *special* features?
- Numeric Results
  - Most probable outcome(s).
  - Finding the Expected Value (or Mean Value) $E(X) = \mu$
  - Determining the Variance Var(X) $= \sigma^2$ and Standard Deviation σ.
  - Formulas:

$$
\begin{aligned}
E(X) &= \sum x \cdot p(x) \\
\mathrm{Var(X)} &= E(X^2) - [E(X)]^2 = \left(\sum x^2 \cdot p(x)\right) - \mu^2
\end{aligned}
$$

Bernoulli Distribution

- Situation: Yes/No
- Parameters: p=Yes, $q = 1 - p$=No
- Outcomes: X= 0 or 1
- Formula: $P(X = 1) = p$, $P(X = 0) = 1 - p$
- Mean: $\mu = p$
- Variance: $\sigma^2 = p(1 - p) = pq$, $\sigma = \sqrt{pq} \leq \frac{1}{2}$

Binomial Distribution

- Situation: Fixed Number of Independent Yes/No Trials
- Parameters: n Trials, p=Yes
- Outcomes: $X = 0, 1, \ldots, n - 1, n$
- Formula: $P(X = k) = \binom{n}{k} p^k q^{n-k}$
- Calculation Formula: $P(X = k + 1) = P(X = k) \cdot \frac{(n-k)p}{(k+1)q}$
- Mean: $\mu = np$
- Variance: $\sigma^2 = npq$, $\sigma = \sqrt{npq}$
- Largest Probability: The smallest value of k so that $k \geq np - q$

Poisson Distribution

- Situation: Counting Events in a Fixed Amount of Time (or Space)
- Parameter: Mean Number of Events λ
- (Infinite) Outcomes: $X = 0, 1, 2, \ldots$
- Formula: $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$
- Calculation Formula: $P(X = k + 1) = P(X = k) \cdot \frac{\lambda}{k+1}$
- Variance: $\sigma^2 = \lambda$, $\sigma = \sqrt{\lambda}$
- Largest Probability*: The largest value of $k \leq \lambda$.

Geometric Distribution

- Situation: Counting Attempts Until (and Including) a Success
- Parameter: Success Probability p
- (Infinite) Outcomes: $X = 1, 2, \ldots$
- Formula: $P(X = k) = p \cdot q^{k-1}$
- Formula #2: $P(X \geq k) = q^{k-1}$
- Calculation Formula: $P(X = k + 1) = P(X = k) \cdot q$
- Mean: $\mu = \frac{1}{p}$
- Variance $\sigma^2 = \frac{q}{p^2}$, $\sigma = \frac{\sqrt{q}}{p}$
- Largest Probability: $X = 1$

Negative Binomial Distribution

- Situation: Count Attempts Until (and Including) a fixed Number of Successes
- Parameters: Success Probability p, Number of Successes r
- (Infinite) Outcomes: $X = r, r + 1, r + 2, \ldots$
- Formula: $P(X = k) = \binom{k-1}{r-1} p^r q^{k-r}$
- Calculation Formula: $P(X = k + 1) = P(X = k) \cdot \frac{kq}{k-r+1}$
- Mean: $\mu = \frac{r}{p}$
- Variance: $\sigma^2 = \frac{rq}{p^2}$, $\sigma = \frac{\sqrt{rq}}{p}$
- Largest Probability: The smallest allowed k so that $k \geq \frac{r-1}{p}$

Uniform Distribution

- Situation: Pick From a Set of Consecutive Integers
- Parameter: Lowest and Highest integers a, b
- Outcomes: $X = a, a+1, a+2, \ldots, b-2, b-1, b$
- Formula: $P(X = k) = \frac{1}{(b-a+1)}$
- Mean: $\mu = \frac{a+b}{2}$
- Variance: $\frac{(b-a+1)^2-1}{12}$
- No Largest Probability

Hypergeometric Distribution

- Situation: Sampling a Fixed Number of Items For Success from a Larger Finite Population
- Parameters: Population Size N, Sample size n, Successful Items in Population M
- Outcomes*: $X = 0, 1, ...n$
- Formula: $P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$
- Calculation Formula?: $P(X = k+1) = P(X = k) \cdot \frac{(M-k)(n-k)}{(k+1)(N-M-n+k+1)}$
- Mean: $\mu = \frac{nM}{N}$ (Same as binomial)
- Variance: $\sigma^2 = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}\left(= npq\frac{N-n}{N-1}\right)$
- Largest Probability?

Multinomial Distribution

- Situation: Tracking More Than Two Outcomes Simultaneously
- Parameters: Number of Attempts n, probabilities p_i, $\sum p_i = 1$
- Outcomes: $(X_1, X_2, X_3, \ldots) = (k_1, k_2, k_3, \ldots)$ $\sum k_i = n$
- Formula: $P(\vec{X} = \vec{k}) = \binom{n}{k_1, k_2, k_3, \ldots} p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots$
- Multinomial Coefficient: $\binom{n}{k_1, k_2, k_3, \ldots} = \frac{n!}{k_1! k_2! k_3! \cdots}$
- Mean: $\mu_i = np_i$
- Variance: $\sigma_i^2 = np_i(1 - p_i)$
- Largest Probability: Near $k_1 = np_1$, $k_2 = np_2$, $k_3 = np_3$, etc.

Normal Distribution

- It's a Continuous Distribution
- Situation: Commonly Occurs and Is A Limit of Discrete Distributions (CLT)
- Parameters: Mean μ and Standard Deviation σ
- Formula: Look Up In A Table (, Calculator or Computer)
- Conversions: Standard Normal ($\mu = 0$, $\sigma = 1$) uses $Z = \frac{X - \mu}{\sigma}$
- Largest (Local) Probability: At The Mean μ

Binomial Coefficients

- What They Mean? $\binom{n}{k}$ "n Choose k": Counting Subsets of a Larger Set
- Binomial Theorem $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$
- Binomial Series $(1+x)^p = \sum_{k=0}^{\infty} \binom{p}{k} x^k$
- Basic Definition (Don't Use If Possible!): $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
- Simple Situations: $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = n$
- Extended Definition: (p not integer) $\binom{p}{k} = \frac{p(p-1)\cdots(p-k+1)}{k!}$

Binomial Identities

- Can you prove the following identities?
- $\binom{n}{k} = \binom{n}{n-k}$
- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$ (Even if k is odd)
- $\sum_{k=0}^{n} 2^k \binom{n}{k} = 3^n$ (Even if k is odd)
- Recurrence Relation #1: $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$
- Recurrence Relation #2: $\binom{n}{k} = \binom{n}{k-1} \cdot \frac{n-k+1}{k}$
- Recurrence Relation #3: $\binom{n}{k} = \binom{n-1}{k} \cdot \frac{n}{n-k}$

Counting Candy

- Number of subsets from n items: 2^n
- Number of k-subsets from n types with repetition: $\binom{n+k-1}{k}$

Independence v. Dependence

- Does Knowledge About One Event A (or variable X) Affect the Probability of Event B (or variable Y)?
- Conditional Probability Notation $P(A|B)$ "probability of A assuming that B is known".
- Formulas:
  - Area (Dartboard) Analogy
  - Division: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
  - Multiplication: $P(A \cap B) = P(A|B)P(B)$
  - Generally $P(A|B) \neq P(A)$
- Independence
  - Basic Idea: Knowledge Has No Effect
  - Precise Definition: $P(A|B) = P(A)$
  - Multiplication Formula $P(A \cap B) = P(A) \cdot P(B)$
- Collections of Random Variables
  - X and Y, Sum $S = X + Y$
  - Always: $E(X + Y) = E(X) + E(Y)$
  - If X, Y Independent, $\mathrm{Var(X + Y) = Var(X) + Var(Y)}$
  - X,Y assumed Independent and Identical: Binomial, Negative Binomial, Poisson?
  - X, Y Identical but not Dependent: Hypergeometric
  - n Identical, Independent X_i: $E(S) = nE(X)$, $\mathrm{Var(S) = nVar(X)}$
  - Dependent: $\mathrm{Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)}$
  - Covariance: $\mathrm{Cov(X, Y) = E(XY) - E(X)E(Y)}$
  - Inequality: $\mathrm{|Cov(X, Y)| \leq \sqrt{Var(X)Var(Y)}}$

Taylor Series for $f(x)$

- Method of approximating a function $f(x)$ near a known value x_o.
- First order approximation (basic calculus): $f(x) = f(x_0) + f'(x_0)(x - x_0)$
- Higher Approximations: $f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots = \sum_{k=0}^{\infty} \frac{f^{(n)}}{n!}(x_0)(x - x_0)^n$.
- In practice: One through Three Derivatives.
- Warnings:
  - Not Exact. Error: $E \approx K|x - x_0|^{n+1}$ where K is derived from $\frac{f^{(n+1)}(x)}{(n+1)!}$.
  - Requires *Nice* (Continuous) Derivatives For $f(x)$.
- Examples:
  - $\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 \cdots$
  - $e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \cdots$
  - $\ln(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots$

Big-Oh & Little-oh Notation

Stirling's Formula

- Method of Approximating $n!$ for Large N by a Simpler Expression.
- Method:
  - Take Logarithms: $\ln n! = \sum_{k=1}^{n} \ln n$
  - Approximate by Integrals: $\int_1^n \ln x \, dx \leq \sum_{k=1}^{n} \ln n \leq \int_2^{n+1} \ln x \, dx$
  - $n \ln n - n + 1 \leq \ln n! \leq (n+1)\ln(n+1) - (n+1) + 2 - 2\ln 2$
  - $e\left(\frac{n}{e}\right)^n \leq n! \leq \frac{e^2}{4}\left(\frac{n+1}{e}\right)^{n+1}$
  - Stirling's Formula: $\boxed{n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n}$

Binomial Approximations

Models: Walks & Urns

Generating Function Methods

Markov Chain Methods