APPENDIX C

PROBABILITY THEORY

…you can never know too much probability theory. If you are well grounded in probability theory, you will find it easy to integrate results from theoretical and applied statistics into the analysis of your applications.—Daniel McFadden, 2000 Nobel Prize in Economics

• Organization of appendix in ISSO
  – Basic properties
    • Sample space
    • Expected value
  – Convergence theory
    • Definitions (four modes)
    • Examples and counterexamples
    • Dominated convergence theorem
    • Convergence in distribution and central limit theorem
Probability Theory

- Random variables, distribution functions, and expectations are critical
  - Central tools in stochastic search, optimization, and Monte Carlo methods
- Probabilistic convergence important in building theoretical foundation for stochastic algorithms
- Most theoretical results for algorithms rely on asymptotic arguments (i.e., convergence)
Expectation

• Let $X \in \mathbb{R}^m$, $m \in \{1, 2, \ldots\}$, be distributed according to density function $p_X(x)$

• Then the expected value of a function $f(X)$ is

$$E[f(X)] = \int_{\mathbb{R}^m} f(x) p_X(x) \, dx$$

provided that $E(\|f(X)\|) < \infty$

• Obvious analogue to above for discrete random vectors

• Important special cases for expected value:
  – Mean: $f(X) = X$
  – Covariance matrix: $f(X) = [X - E(X)] [X - E(X)]^T$
Probabilistic Convergence

- Finite-sample results are usually hopeless
- Asymptotic (convergence) results provide means to analyze stochastic algorithms
- Four famous modes of convergence:
  - almost surely (a.s.)
  - in probability (pr.)
  - in mean-square (m.s.)
  - in distribution (dist.)
- First three modes above pertain to sense in which $X_k \rightarrow X$ as $k \rightarrow \infty$
- Last mode (dist.) pertains to convergence of distribution function of $X_k$ to distribution function of $X$
Implications for Four Modes of Convergence

pr.  a.s.

m.s.  dist.