Statistics 430
Summary, Spring 2003

Probability

- Counting methods, sample space S, events; $P(E) = \#E/\#S$
- Unions $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Conditional probability $P(A|B) = P(A \cap B)/P(B)$
- Intersections $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$
- Complements $P(A^c) = 1 - P(A)$
- Law of total probability $P(A) = \sum_j P(A|B_j)P(B_j), \quad \cup B_j = B$
- Bayes $P(A|B) = P(B|A)P(A)/P(B)$
- Independent vs. dependent events

Random variables

- Properties of PDF and CDF
- Mean, variance, and standard deviation of random variable
- Marginal, conditional, joint distributions
- Relationships that connect the different types
  (*e.g.*, Poisson process and exponential r.v.)
- Independence of random variables; sums of random variables
- Markov chain (sequence of random variable)

		Mean	Variance
Bernoulli	$p(x) = p^x(1-p)^{1-x},\ x = 0, 1$	p	$p(1-p)$
Binomial	$p(x) = \binom{n}{x} p^x(1-p)^{n-x},\ x = 0, \ldots, n$	np	$n\,p\,(1-p)$
Geometric	$p(x) = (1-p)^{x-1}p,\ x = 1, 2, \ldots$	$1/p$	$(1-p)/p^2$
Poisson	$p(x) = e^{-\lambda}\lambda^x/x!,\ x = 0, 1, \ldots$	λ	λ
Uniform	$f(x) = 1, \quad 0 \le x \le 1$	$1/2$	$1/12$
Exponential	$f(x) = \lambda e^{-\lambda x},\ x \ge 0$	$1/\lambda$	$1/\lambda^2$
Normal	$f(x) = e^{-(x-\mu)^2/(2\sigma^2)}/\sqrt{2\pi\sigma^2}, -\infty < x < \infty$	μ	σ^2

Expected value

- Definition of $E\,X$ as a weighted sum, with weights given by $p(x)$.
- Expected value of a sum = sum of expected values
  $E(a + bX + cY) = a + b\,E(X) + c\,E(Y)$
- Covariance, correlation and variance of sums; portfolios
  $\mathrm{Cov}(X, Y) = E((X - EX)(Y - EY)) = E(XY) - (EX)(EY)$
- Variance of a sum = sum of variances *if* all covariances are zero, otherwise
  $\mathrm{Var}(aX + bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y)$
- Expected value of indicator variable, $E\,I = P(I = 1)$
- Conditional expected value, $E(X|Y)$; use in prediction
- Marginal from conditional $E\,X = \sum_y E(X|Y = y)P(Y = y)$

Inequalities and asymptotics

- Bonferroni $P(\cup_i A_i) \leq \sum_i P(A_i)$
- Jensen $f(E(X)) \leq E(f(x))$ if $f(x)$ is convex (*e.g.*, $f(x) = x^2$)
- Markov $P(X \geq k) \leq EX/k$ if X is a non-negative r.v.
- Chebyshev $P(|X - EX|/SD(X) \geq k) \leq 1/k^2$
- Central limit theorem $\sum_i X_i \to$ Normal
- Weak law of large numbers; averages get close their expected value