MATHEMATICS 191, FALL 2004
MATHEMATICAL PROBABILITY

Last revised: September 14, 2004

Instructor: Paul Bamberg
Offices: SC 423, 495-1748 and Quincy House 102, 493-3100. Quincy 102 opens off the Quincy House courtyard, near the raised cubical library.
Email: `bamberg@tiac.net`

Office Hours:

- Tu 2:30-3:30 in Science Center 423
- Tu 10 AM - noon, W 9AM - 2PM, and Th 8 AM - noon in Quincy 102 (phone 3-3100 first)
- Tu evenings in Quincy 102, but phone 3-3100 first)

Course Assistant: Mike Chiappa, `chiappa@fas.harvard.edu`
Course Website: `http://www.courses.fas.harvard.edu/-math191` (That's a tilde before math191)

Goals and Prerequisites: This is a moderately rigorous course in probability theory and its applications. A nice feature of probability is that it presents subtle issues involving the infinite in contexts that appear realistic in the sense that you could imagine betting on the outcome of the questions that was posed. So while the course will require only infinite series (Math 1b), multivariable calculus (Math 21a) and a bit of linear algebra (Math 21b), it will lead naturally into areas such as set theory and elementary measure theory.

One freshman took the course last year and did well. Any freshman who did well on the online placement test that covers Math 21a material is welcome to take the course.

Placement Tests: This year there are some experimental online math placement tests designed to help freshmen who are considering taking math courses numbered higher than 21. We need to know how students like you, who are clearly qualified for such courses, score on these tests, so that next year we can give accurate placement advice based on the tests.

The tests will only be online for the first couple of weeks of the term. During that period, please log on to `math.placement.fas.harvard.edu` and take the tests called "Math 21a-mastery" and "more fun math." Doing this will be a public service and will also help your grade slightly. To get credit, just send an email to `bamberg@tiac.net` with the information from the "receipts" that are created when you finish the test.

Course Meetings: The course meets TTh from 4-5:30 P. M. in Science Center 110. This strange time, which conflicts with almost nothing because voting members of the Faculty are not allowed to teach then, was chosen in response to a request from a Physics concentrator to break the time conflict with introductory physics courses. This is important math for physicists to know, and I hope that physics concentrators will take advantage of the opportunity.

There will be additional weekly problem sessions led by the course assistant(s). These will not be the usual optional "help sessions." Instead, members of the class will be assigned problems to present at the blackboard, drawn from the collection of 1000 solved problems that accompanies the textbook. Participation in these sessions will count for one-third of your homework grade, and graded written problem sets will be shorter than usual.

Grades: Your course grade will be determined as follows:

- required homework, 40 points
- problem session presentations, 20 points
- two best quizzes, 20 points each
- third quiz, 10 points
- final exam, 100 points
- 2 placement tests, 10 points each (also added to the total possible points)

The total points available are thus 210 or 230, depending on whether you do the placement tests. The grading scheme is as follows:

Points	Minimum Grade
92%	A
86%	A-
80%	B+
74%	B
68%	B-
62%	C+
56%	C

I reserve the right to be more generous if exams prove unexpectedly long or difficult.

Exams: There will be three in-class quizzes and one final exam. The quizzes will be roughly one-half hour each, and the final is scheduled for three hours. There will be an opportunity to retake each quiz, but the maximum possible score from a makeup quiz is 80

Three Quizzes:	Thursday, October 14
	Thursday, November 18
	Thursday, December 16
Final Exam:	comprehensive, with fairly uniform coverage

Texts:

Probability and Random Processes, Grimmett and Stirzaker, third edition, Oxford University Press, 2002, ISBN# 0-19-857222-0 (at the Coop)

One Thousand Exercises in Probability, Grimmett and Stirzaker, Oxford University Press, 2002, ISBN# 0-19-857221-2 (at the Coop)

Elementary Probability, Stirzaker, Second Edition, Cambridge University Press, 2003, ISBN# 0-521-53428-3 (at the Coop)

The course will follow the order of topics in the first book. The third book (Stirzaker) covers most of these topics in a somewhat different order and in more detail, but it omits some of the most interesting material and does not take full advantage of the mathematical background and sophistication of Harvard students.

Homework and Programming Assignments: Homework (five or six problems) will be assigned weekly and will be due at the start of Thursday's class. The CA will return your corrected homework to you at the following class.

Six problems will be assigned for each weekly section. You will be responsible for preparing three of these to present at the blackboard, and exams will assume that you become familiar with all six. The solutions are all in One Thousand Exercises in Probability, but they take a while to understand!

You are encouraged to discuss the course with other students, your CA and the instructors, *but you should always write your homework solutions out yourself in your own words.*

List of Topics(dates are not accurate in this draft):

Date		Chapter	Topics
September	21	Apostol	Countability and Uncountability
	23	1	Sets and Probability
	28	1	Combinatorics, Bridge, Poker, and Dice
	30	1	Conditional Probability and Independence
October	5	1	Simpson's Paradox explained by Conditional Probability
	7	2	Random Variables, Law of Averages
	12	3	Discrete Distributions and Random Variables
	14	3	Expectation and Variance - QUIZ 1
	19	3	Matching Problems
	21	3	Poisson and Negative Binomial Distributions
	26	3	Simple Random Walk
	28	3	Ballot theorem and Reflection Principle
November	2	3	Arc Sine Laws for Random Walks
	4	4	Continuous Distributions and Random Variables
	9	4	Functions of Random Variables
	16	4	Multivariate Normal Distribution
	18	4	2 Random Variables - QUIZ 2
	23	4	Change of variable; convolution
	30	4	Geometric Probability - Bertrand's Paradox
December	2	4	Geometric Probability - Crofton's method
	7	5	Generating Functions
	9	5	Random Walk done via Generating Functions
	14	5	Weak Law of Large Numbers, Central Limit Theorem
	16	7	Strong Law of Large Numbers - QUIZ 3
	21	6	Linear Algebra and Markov Chains

Chapters 1-5 of Grimmett and Stirzaker will be covered quite thoroughly. Coverage of chapters and 7 will be limited to a couple of interesting topics.

1 Checklist of concepts, applications, and theorems

This is what you should be able to do at the end of the course.

1.1 Definitions and simple theorems

1. Countability - finite Cartesian products, countable unions and arbitrary intersections of countable sets are countable. The collection of all subsets of a countably infinite subset is uncountable.
2. Definition of a σ-field - closure under complement and countable union implies closure under difference and intersection. Example of two σ-fields whose union is not a σ-field.
3. Inclusion-exclusion formula for two events, and related inequalities.
4. Probabilities for increasing and decreasing infinite sequences of events.
5. Definition of conditional probability.
6. Definition of independent events; example that three events can be pairwise independent but not independent.
7. Probability of k successes in n independent identical experiments.
8. Geometric distribution - probability that the first success in a sequence of identical independent experiments occurs on the kth attempt.
9. Definition of random variables - discrete, continuous, and those that are neither.
10. General properties of distribution functions for random variables.
11. Definition of indicator functions.
12. Definition of independence for discrete random variables.
13. Probability mass functions for binomial, Poisson, geometric and negative binomial distributions.
14. Definition of expectation and variance for discrete random variables; examples where these are not defined because the series is not absolutely convergent.
15. Definition of "uncorrelated" for two random variables; example of two random variables that are uncorrelated but not independent.
16. Statement and proof of the "law of the unconscious statistician" for one discrete random variable.

17. Expectation of the sum of random variables.

18. Variance of the sum of independent random variables and of a multiple of a random variable.

19. Definition of covariance and correlation for two random variables.

20. Basic properties of random walks - homogeneity in time and space. Markov property.

21. Random walks - formula for number of paths in n steps from level a to level b.

22. Reflection principle for random walks - use to obtain a formula for the number of paths from level a to level b that do not cross level 0.

23. Ballot theorem, including the equivalent case of the "hitting time theorem:" the fraction of paths from level 0 to level b in N steps that reach level b for the first time at the Nth step.

24. Key lemma for arc sine laws. For a symmetric random walk ($p = \frac{1}{2}$) that starts in level 0, the following are all equal:

    - u_{2m} = the probability of being in level 0 after $2m$ steps.
    - u_{2m} = the probability of never revisiting level 0 during $2m$ steps.
    - u_{2m} = the probability of never visiting a negative level during $2m$ steps.
    - u_{2m} = the probability of never visiting a positive level during $2m$ steps.

25. By Sterling's approximation, for a symmetric random walk,

$$u_{2k} = 2^{-2k} \frac{(2k)!}{(k!)^2}$$

    is approximately $\frac{1}{\pi k}$

26. Arc sine law for last return (mass function): after $2n$ steps of a symmetric random walk, the probability that the last visit to level 0 was at step $2k$ is exactly $u_{2k}u_{2n-2k}$ and approximately

$$\frac{1}{\pi\sqrt{k(n-k)}}$$

27. Arc sine law for last return (distribution function): after $2n$ steps of a symmetric random walk, the probability that the last visit to level 0 was before step $2xn$ is approximately $\frac{2}{\pi} \arcsin \sqrt{x}$.

28. Arc sine law for sojourn times (mass function): after $2n$ steps of a symmetric random walk, the probability that $2k$ segments were in "positive territory" is exactly $u_{2k}u_{2n-2k}$ and approximately

$$\frac{1}{\pi\sqrt{k(n-k)}}$$

29. Arc sine law for sojourn times (distribution function): after $2n$ steps of a symmetric random walk, the probability that fraction of time spent in "positive territory" is approximately $\frac{2}{\pi}\arcsin\sqrt{x}$.

30. General properties of density functions for continuous random variables.

31. In terms of the density function $f_X(x)$, integrals for expectations $\mathbb{E}(X)$ and $\mathbb{E}(X^2)$

32. "Law of the unconscious statistician" for the expectation of a function $Y = g(X)$ of a continuous random variable.

33. Exponential distribution and its expectation.

34. Density function for normal distribution $N(0,1)$ - the variance is 1.

35. Density function for normal distribution $N(\mu,\sigma)$ - the mean is μ and the variance is σ.

36. For a pair of random variables, express the distribution function $F_{X,Y}(x,y)$ in terms of the density function $f_{X,Y}(x,y)$ and vice versa.

37. "Law of the unconscious statistician" for calculating the expectation of a random variable $Z = g(X,Y)$.

38. Definition of marginal density and conditional expectation.

39. Definition of independence for two random variables; what this means in terms of the density function.

40. Density function for the sum of two independent random variables X and Y as a convolution of the density functions for X and Y.

41. Bertrand's paradox: state the condition for an equilateral triangle based on a random chord to fit inside a circle. Describe three reasonable but different interpretations of "random chord," and for each of them determine the probability that the triangle fits inside the circle.

42. Crofton's method: the general strategy for turning a geometrical probability problem into a differential equation.

43. Generating functions: definition of $G(s)$ in terms of the probability mass function. Determining a value of the probability mass function by evaluating an appropriate derivative of $G(s)$. Significance of $G(1)$.

44. Generating function for the sum of two or more independent random variables or for the sum of n independent copies of a random variable.

45. Statement of the "weak law of large numbers."

46. Statement of the central limit theorem, and the recipe that it implies for generating random numbers with a normal distribution.

47. Statement of Markov's inequality and Chebyshov's inequality, and use to obtain crude bounds on how fast the probability must fall to zero for large values of a random variable with zero mean.

48. Statement of the "strong law of large numbers" and what makes it "stronger." than the weak law.

1.2 Applications

1. Equations and inequalities that follow by using induction and the inclusion-exclusion formula.

2. Calculating probabilities by using the inclusion-exclusion formula for two or three events.

3. Probabilities for poker hands.

4. Probabilities for bridge - distribution of suits, or distribution of missing cards in a suit between two opposing hands.

5. Conditional probability problems involving two events (the "bearded man" problem).

6. Generalizations of the Monty Hall problem.

7. Problems based on "Bernoulli trials" - a certain number of successes in n independent identical experiments.

8. Conditional probability problems based on the veracity of witnesses (Eddington's controversy, Queen of Sheba)

9. Analysis of Simpson's paradox in terms of conditional probability.

10. Problems based on geometric and negative binomial distributions - what is the probability that it takes exactly n trials to get k successes?

11. "Problem of the points" - what is the probability that m successes occur before n failures?

12. Problems that involve drawing (without replacement) from an urn containing r red balls and b blue balls.

13. Analysis of union, intersection, etc. in terms of indicator functions.

14. Problems involving the "probabilistic method" and the pigeonhole principle.

15. Problems based on the Poisson distribution.

16. Random walk problems, solved by using conditional probability to write down a linear difference equation and then solving the difference equation with appropriate boundary conditions.

17. Apply the key lemma for arc sine theorems to get various results about random walks (involving maximum or minimum level, for example)

18. For random walks with a small even number of steps, get exact results for the probability that the last visit to the origin was at some specific time $2r$.

19. Given the density or distribution function for a random variable X, determine its expectation and variance.

20. Given the density or distribution function for a random variable X and a new random variable $Y = g(X)$, determine the density function and expectation of Y.

21. Given a random variable X with a uniform distribution on [0,1], invent a random variable $Y = g(X)$ that has some specified distribution function or density function.

22. Given a joint density function $f_{X,Y}(x, y)$, determine whether or not X and Y are independent.

23. Given a joint density function $f_{X,Y}(x, y)$, calculate the conditional density function $f_{Y|X}(y, x)$.

24. Given a joint density function $f_{X,Y}(x, y)$ and new random variables U and V defined by formulas $u(x, y)$ and $v(x, y)$, determine the density function $f_{U,V}(u, v)$

25. Given the density functions for two independent random variables, calculate the density function for their sum.

26. Given a specification of a way to select the "random chord" for Bertrand's paradox, calculate the joint density function for the polar coordinates of the midpoint of the chord.

27. Solve geometric probability problems involving two points by dividing the region of interest in half and conditioning on whether the two points lie in the same half or in different halves.

28. Solve geometric probability problems where one of the points of interesting lies at a randomly chosen point on a line or circular arc by conditioning on the position of that point.

29. Solve geometric probability problems by Crofton's method, considering first the simpler case where a point of interest lies on the boundary, then setting up a differential equation.

30. Write down the generating function for a discrete random variable that can take a small number of integer values and for the sum of several such variables.

31. Given the generating function for a random variable X, determine the expectation and variance of X.

32. Given a generating function for the sum S of k independent copies of a random variable, expand in a power series to find the probability that $S = n$. For example, what is the probability to throw a total of n with k fair dice?

33. Use generating functions to solve problems involving the sum of random variables with the geometric distribution.

34. Given a random variable X (discrete or continuous), determine what the law of large numbers has to say about the average of many such random variables.

35. Given a random variable X (discrete or continuous), calculate its mean and variance, and construct a rescaled sum of n independent copies of X whose distribution, according to the central limit theorem, will be approximately normal.

36. Analyze a 2-state or 3-state Markov chain in terms of eigenvalues and eigenvectors of a matrix.

1.3 Proofs

1. Bernstein's inequality and the law of averages: Suppose that the probability of a head when a coin is tossed is p. Define the random variable S_n to be the number of heads in n tosses. Prove that the probability

$$\mathbb{P}(\frac{S_n}{n} \geq p + \epsilon) \leq e^{-\frac{1}{4}n\epsilon^2}$$

2. Given the joint distribution function for two discrete random variables X and Y, derive a formula for $\mathbb{P}(X = x, Y = y)$.

3. Given the joint distribution function for two continuous random variables X and Y, derive a formula for $\mathbb{P}(a < x < leqb, c < Y \leq d)$.

4. Calculate the mean and variance for the binomial distribution.

5. Obtain the Poisson distribution as a limiting case of the binomial distribution, and calculate its mean and variance.

6. Use indicator functions to prove that if n letters are placed randomly into n matching envelopes, the probability that exactly r letters are placed in the correct envelope is

$$\mathbb{P}(X = r) = \frac{1}{r!} \sum_{s=0}^{n-r} \frac{1}{s!}$$

7. Prove the Cauchy-Schwarz inequality for the case of two random variables with zero mean, and use it to show that correlation cannot exceed 1 in magnitude.

8. Gambler's ruin: prove that for a symmetric simple random walk, starting in level k with an absorbing barrier at level 0, the probability of reaching level 0 is 1 but the expected time to do so is infinite.

9. Prove the ballot theorem from the reflection principle.

10. Either by using a combinatorial argument based on the ballot theorem or by a geometrical argument establishing a bijection between paths, prove that the number of paths that start in level 0 and are again in level 0 after $2m$ steps is equal to the number of paths that do not return to level 0 during $2m$ steps.

11. By a geometrical argument establishing a bijection between paths, prove that the number of paths that start in level 0 and do not return to level 0 during $2m$ steps is equal to the number of paths that start in level 0 and do not visit any negative level during $2m$ steps.

12. Prove that after $2n$ steps of a symmetric random walk, the probability that the last visit to level 0 was at step $2k$ is exactly $u_{2k}u_{2n-2k}$.

13. Prove that after $2n$ steps of a symmetric random walk, the probability that the last visit to level 0 was before step xn is approximately $\frac{2}{\pi} \arcsin \sqrt{x}$.

14. Prove that the probability f_{2r} that a symmetric random walk that started in level 0 is again in level 0 for the first time after $2r$ steps is $f_{2r} = \frac{1}{2r-1} \frac{(2r)!}{(r!)^2}$.

15. By conditioning on the time $2r$ of first return to the origin, prove that after $2n$ steps of a symmetric random walk, the probability that $2k$ segments were in "positive territory" is exactly $u_{2k}u_{2n-2k}$.

16. Prove that for a random variable with a density function $f_X(x)$ that is 0 for negative x,

$$\int_0^\infty \mathbb{P}(X > x)dx = \mathbb{E}(X)$$

.

17. For the special case of a random variable $Y = g(X)$ where $g(x) > 0$ for all x, prove the law of the unconscious statistician,

$$\mathbb{E}(Y) = \int_0^\infty g(x)f_X(x)dx$$

.

18. Given the density function for normal distribution

$N(\mu, \sigma)$,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

make a change of variable and invoke properties of the distribution $N(0, 1)$ to show that the mean is μ and the variance is σ^2.

19. Prove that for the bivariate normal distribution

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{(x^2-2\rho xy+y^2}{2(1-\rho^2)}},$$

var(X) = var$(Y) = 1$ and cov$(X,Y) = \rho$.

20. Prove the convolution formula for the density function of the sum of two continuous random variables $Z = X + Y$.

21. Buffon's needle: Consider a fixed needle of length L whose midpoint is at the center of a circle of diameter d. Choose a random diameter of the circle, then construct a chord whose midpoint lies at a point chose with uniform density on the diameter. Prove that the probability that the needle intersects the chord is $\frac{2L}{\pi d}$.

22. By interchanging the order of summation over two indices, prove that the product of the generating functions for two discrete random variables is the generating function for the convolution of their mass functions and is therefore the generating function for their sum.

23. Derive the formula for the variance of a random variable in terms of its generating function.

24. Use generating functions to calculate mean and variance for the Poisson distribution, show that the sum of two independent Poisson random variables is also Poisson.

25. Let p_n be the probability that when n letters are stuffed randomly into n matching envelopes, no letter ends up in its correct envelope. Starting from the difference equation, valid for $n \geq 3$,

$$np_n = p_{n-2} + (n-1)p_{n-1},$$

derive and solve a differential equation to show that

$$G(s) = \sum_{n=1}^{\infty} p_n s^n = \frac{e^{-s}}{1-s} - 1$$

26. Use generating functions to prove Waring's theorem: Given events A_1 $A_2 ... A_n$ with

$$S_m = \sum_{i_1 < i_2 < ... < i_m} \mathbb{P}(A_{i_1} \cap A_{i_2} ... \cap A_{i_m}),$$

the probability that precisely X of the events occur is

$$\mathbb{P}(X = i) = \sum_{j=i}^{n} (-1)^{j-i} \frac{j!}{(j-i)!i!} S_j$$

27. Let $f_r(n)$, for $r > 0$, be the probability that a simple random walk (up one level with probability p, down one level with probability $q = 1 - p$) reaches level r for the first time at the nth step. Define the generating function

$$F_r(n) = \sum_{n=r}^{\infty} f_r(n) s^n.$$

Explain why $F_r(s) = [F_1(s)]^r$, and show that

$$F_1(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}$$

28. Suppose that independent random variables X_i all have the same distribution function, and expectation $\mathbb{E}(X_i) = \mu$. Let $S_n = \sum_{i=1}^{n} X_i$. Prove that as $n \to \infty$ the generating function for $\frac{S_n}{n}$ approaches the constant μ.

29. Suppose that independent random variables X_i all have the same distribution function, with expectation 0 and variance $\mathbb{E}(X_i^2) = \sigma^2$. Let $S_n = \sum_{i=1}^{n} X_i$. Prove that as $n \to \infty$ the generating function for $\frac{S_n}{n}$ approaches the generating function for the normal distribution $N(0, 1)$.

30. State and prove the strong law of large numbers for the case of a random variable with zero mean for which $\mathbb{E}(X^4)$ is finite.