Math 510 Final Review Sheet

1 Topics

1. Combinatorics
   (a) Basic principle of counting: multiplication
   (b) repeated choices from a set: with or without order, with or without repetition
      • with order, with replacement: \( n^r \)
      • with order, without replacement: permutations: \( n!/(n-r)! \)
      • number of orders of a set: \( n! \)
      • When some objects are repeated: multinomial coefficients
      • without order, without replacement: combinations: \( \binom{n}{r} = \frac{n!}{r!(n-r)!} \)
      • without order, with replacement: \( \binom{n+r-1}{r-1} \)
      • without order, with replacement, at least one each: \( \binom{n-1}{r-1} \)
      • multinomial coefficients
   (c) combinations, binomial theorem, Pascal’s triangle
   (d) multinomial theorem

2. Probability
   (a) Setup: outcomes, sample space, events, set theory
   (b) How an event is both “something that can happen or not” and “a set of outcomes”
   (c) Axioms for probability
   (d) Some propositions for probability
   (e) Solving problems by considering the complement
   (f) Inclusion-Exclusion principle
   (g) Equally likely outcomes
   (h) various interpretations of probability: repeated trials, subjective feeling

3. Conditional Probability
   (a) \( P(E|F) = \frac{P(EF)}{P(F)} \)
   (b) Reduced sample space
   (c) Multiplication rule and tree diagrams
   (d) Bayes’ formula
(e) odds

(f) independent events
   • when we should expect events to be independent
   • \( P(E|F) = P(E); P(EF) = P(E)P(F) \), symmetry with respect to interchanging \( E \) and \( F \)
   • consequences for complements
   • Series of many independent events

(g) \( P(\cdot|F) \) as a probability

4. Random Variables (general)

(a) Definition

(b) Cumulative distribution function
   • definition: \( P(X \leq t) \)
   • Using it to find probabilities of events
   • Properties (see below)
   • Using it to relate \( X \) to \( g(X) \)

(c) Expected value
   • definition for discrete
   • definition for continuous
   • intuitive notion of average (i.e. mean)
   • Expected value of \( g(X) \)
   • \( E[X + a], E[cX], \) and using these facts

(d) Variance, standard deviation
   • definition as \( E[(X - \mu)^2] \)
   • alternate formula: \( E[X^2] - E[X]^2 \)
   • Standard deviation as \( \sqrt{Var(X)} \)
   • intuitive notion of measuring “average” distance from mean
   • non-negative; positive unless constant
   • \( Var(X + a), Var(cX), \) and using these facts
   • Same for standard deviation

5. Discrete random variables: the mass distribution function
   • and relation to cumulative distribution function
   • Using mass function to find probabilities of events
   • Properties: all \( \geq 0 \), sum to 1

6. Continuous random variables: the probability density function
   • and relation to cumulative distribution function
   • Using density function to find probabilities of events
• Properties: always $\geq 0$, integrates to 1

7. Jointly distributed Random Variables
   (a) Using joint mass function (discrete)
   (b) Using joint density function (continuous)
   (c) Marginal distribution functions (mass for discrete; density for continuous)
   (d) Cumulative distribution function
   (e) Independent random variables
      i. Definition
      ii. Factorization of density/mass function as $f(x, y) = g(x)h(y)$
      iii. Sums of independent random variables: convolution

8. Properties of Expected value
   (a) Calculating $E[g(X, Y)]$ by summing over values of $(X, Y)$
   (b) $E[X + Y]$
   (c) $E[XY]$ when $X, Y$ independent

9. Covariance and variance
   (a) Definition of covariance
   (b) Properties of covariance
   (c) Correlation
      • between $-1$ and 1
      • meaning of being 1, $-1$, 0

10. Splitting a random variable as a sum
    (a) Especially for random variables that count something
    (b) Expected values of sum
    (c) Variance of a sum
       • If independent, then add
       • If not, use covariance

11. Conditioning
    (a) conditional mass function (discrete)
    (b) conditional density function (continuous)
    (c) conditional expected value
    (d) conditional variance
    (e) Calculating $E[X]$ by $E[E[X|Y]]$
(f) Calculating $\text{Var}(X)$ by $E[\text{Var}(X|Y)] + \text{Var}(E[X|Y])$

12. Moment generating functions
   (a) Calculating using $E[e^{tX}]$
   (b) Using to find moments
   (c) Sums of independent random variables

13. Inequalities
   (a) Markov's inequality
   (b) Chebyshev's inequality

14. Limit theorems
   (a) What happens to $X_1 + \cdots + X_n$
   (b) What happens to $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$
   (c) What happens to $\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$
   (d) Law of large numbers
      i. probability vs. average
      ii. strong vs. weak
   (e) Central limit theorem

15. Processes
   (a) Bernoulli process
   (b) Poisson process
   (c) Brownian process, modeling the value of an investment using a Brownian process
   (d) Markov chains, transition matrix, limiting distribution

16. Simulation
   (a) Quasirandom number generators
   (b) Discrete distributions
   (c) Continuous distributions using inverse transformation

2 Some predictable kinds of problems

1. Combinatorics: find number of ways something can be done
2. Properties of binomial coefficients
3. Theorems to know, axioms to know, proofs to know
4. Finding a probability, given other probabilities
   - using inclusion-exclusion principle like formulas
   - using Venn diagrams
   - all outcomes equally likely: Same as combinatorics, but with probabilities
   - conditional probability, Bayes’ formula, tree diagrams
   - independence
5. Determining whether something is a mass/density function, or scaling it so that it is one
6. Finding probabilities, from a mass/density function
7. Finding expected value, variance, standard deviation given a mass/density function
8. Recognizing and using a random variable from the list we are supposed to know (above)
9. Finding the expected value of \( g(X) \)
10. Finding a mass/density function for \( g(X) \)
11. Using a table (will be provided) for the normal random variable
12. Using joint mass or density functions to find probabilities
13. Using joint mass or density functions to find expected values of functions \( g(X, Y) \), and covariance, correlation
14. Finding the expected value and variance of a random variable which should be split as a sum of simpler random variables
15. Finding the density function of a sum of two independent random variables
16. Calculating a probability, expected value, or variance by conditioning
17. Calculating a moment generating function
18. Using a moment generating function to find a certain moment, or a variance
19. Applying the Markov or Chebyshev inequality
20. Applying the Central Limit Theorem
21. Dealing with a Poisson process, a Brownian process, or a Markov chain
22. Building a random variable of a given distribution given a uniform \([0, 1]\) random variable (simulation)
3 Particular distributions you should know

- Discrete
  1. Bernoulli (Indicator variables)
  2. Binomial
  3. Poisson (also as a limit of binomial)
  4. Geometric
  5. Hypergeometric

- Continuous
  1. uniform
  2. normal (also as a limit of binomial)
  3. exponential

3.1 What you should about each:

1. Their mass function (for discrete) or density function (for continuous)
2. Recognizing when they are applicable
3. The assumptions inherent in each model
4. What $X$ represents in each model
5. expected value, variance formulas

3.2 Relationships

<table>
<thead>
<tr>
<th>Bernoulli process:</th>
<th>Binomial</th>
<th>Geometric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson process:</td>
<td>Poisson</td>
<td>Exponential</td>
</tr>
</tbody>
</table>

Sums:
- Binomial $(n_1, p)$ + Binomial $(n_2, p)$ is binomial $(n_1 + n_2, p)$
- Poisson $\lambda_1$ + Poisson $\lambda_2$ is Poisson $\lambda_1 + \lambda_2$
- Normal + normal is normal ($\mu$ adds, $\sigma^2$ adds)
- All assuming independence

Limit binomial $\to$ Poisson (when $n$ large, $p$ small, so that $np$ is moderate in size)

Limits that go to normal by central limit theorem:
- binomial when $n$ large
- Poisson when $\lambda$ large
- Sums of i.i.d. random variables
4 Theoretical things to know

4.1 Definitions
- Axioms of probability
- Conditional probability definition \( P(E|F) = \frac{P(EF)}{P(F)} \)
- independence: \( P(E|F) = P(E) \)
- random variable
- mass function
- density function
- cumulative distribution function
- expected value (for discrete)
- expected value (for continuous)
- variance: \( E[(X - \mu)^2] \)
- standard deviation
- independence of random variables
- convolution
- covariance
- correlation
- \( n \)th moments \( \mu_n = E[X^n] \)
- moment generating function \( M_X(t) = \sum_{n=0}^{\infty} \frac{\mu_n}{n!} t^n \)
- Markov chain

4.2 Derivations of formulas
1. number of permutations
2. number of combinations
3. without order, with replacement
4. multinomial coefficients
5. equally likely outcomes: \( P(E) = |E|/|S| \)
6. Bayes' formula
7. \( P(E \cup F) = P(E) + P(F) - P(EF) \)

8. Inclusion-exclusion principle for probability

9. Formulas for mass functions for Bernoulli, binomial, geometric

10. Formulas for \( E[X], Var(X) \) for Bernoulli, binomial, Poisson, geometric, uniform, exponential

11. Alternate formulas for defined concepts (not the definition)
   - Independence: \( P(EF) = P(E)P(F) \)
   - \( Var(X) = E[X^2] - E[X]^2 \)
   - \( Cov(X, Y) = E[XY] - E[X]E[Y] \)
   - \( M_X(t) = E[e^{tX}] \)

12. Conditioning
   - \( P(E) = P(F_1)P(E|F_1) + \cdots + P(F_n)P(E|F_n) \)
   - \( E[X] = E[E[X|Y]] \)
   - \( Var(X) = Var(E[X|Y]) + E[Var(X|Y)] \)

4.3 Theorems to know the statement and proof of

1. Immediate consequences of the axioms of probability
   - \( P(\emptyset) = 0 \)
   - finite versions of axiom 3
   - \( P(E^c) = 1 - P(E) \)
   - If \( E \subset F \), then \( P(E) \leq P(F) \)

2. Properties of mass functions
   - all \( \geq 0 \)
   - Sum to 1

3. Properties of density functions
   - always \( \geq 0 \)
   - Integrates to 1

4. Properties of cumulative distribution functions
   - monotonic increasing
   - \( \lim_{t \to -\infty} F(t) = 0 \)
   - \( \lim_{t \to \infty} F(t) = 1 \)
   - Continuous from the right
• Discrete case: only increases at “jumps”

• Continuous case: Never “jumps”

5. Properties of expected value

• $E[aX] = aE[X]$; $E[X + b] = E[X] + b$ for continuous and discrete

• $X \leq Y$ implies $E[X] \leq E[Y]$

• $E[X + Y] = E[X] + E[Y]$

• When $X$ and $Y$ uncorrelated, $E[XY] = E[X]E[Y]$

6. Properties of variance

• $Var(aX) = a^2Var(X)$; $Var(X + b) = Var(b)$

• When $X$ and $Y$ uncorrelated, $Var(X + Y) = Var(X) + Var(Y)$

7. Properties of Covariance:

• $Cov(X, X) = Var(X)$

• $Cov(aX, Y) = a Cov(X, Y)$

• $Cov(Y, X) = Cov(X, Y)$

• $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$

• $-1 \leq \rho(X, Y) \leq 1$

8. When $X, Y$ independent, then the following equivalent conditions occur: (Note: These do not imply $X$ and $Y$ are independent)

• $E[XY] = E[X]E[Y]$

• $Cov(X, Y) = 0$

• $\rho(X, Y) = 0$

• $X$ and $Y$ are uncorrelated

• $Var(X + Y) = Var(X) + Var(Y)$

9. Properties of moment generating functions

• If $X$ and $Y$ are independent then $M_{X+Y}(t) = M_X(t)M_Y(t)$

• $M_{aX}(t) = M_X(at)$

• $M_{X+b}(t) = e^{bt}M_X(t)$

• Extracting moments from $M(t)$ by examining the series

• Extracting moments from $M(t)$ by differentiation

10. Dealing with $X_1 + \cdots + X_n$ or $\frac{X_1 + \cdots + X_n}{n}$ (assuming i.i.d)

• expected value

• variance or standard deviation
• Central Limit Theorem says as $n$ is large, this is approximately normal

11. Markov’s inequality
12. Chebyshev’s inequality
13. Equivalence of probability and average forms of law of large numbers
14. Weak law of large numbers
15. Strong law of large numbers
16. Central limit theorem

4.4 Theorems to know the statement of
1. binomial theorem/multinomial theorem
2. $E[g(x)]$ in both discrete, continuous case
3. The moment generating function uniquely determines the distribution
4. Unique convergence of Markov chains whose transition matrix has no zeros