UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Midterm 1 Study Guide
Spring 2020

This is a brief summary of the topics we covered that will be in scope for Midterm 1. However, the scope of the exam will encompass all of the homeworks, discussions, lab, and lecture material; if something is not in this document, it is not necessarily out of scope. Students are expected to understand topics in more depth than they are discussed here.

1 Probability Fundamentals

1. Kolmogorov's axioms: probabilities are nonnegative, probability of at least one possible outcome is 1, for disjoint events A, B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
2. Conditional probability: $\mathbb{P}(A, B) = \mathbb{P}(A|B)\mathbb{P}(B)$
3. Law of total probability: for a partition of the sample space (set of disjoint events whose union is the whole sample space) $B_1, B_2, ..., B_n$, $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$
4. Bayes' rule: $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$, derived from $\mathbb{P}(A, B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$
5. Independence implies uncorrelated; reverse not necessarily true
   (a) Independence: A is independent of B if $\forall a, b$, $\mathbb{P}(X = a \cap Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b)$
   (b) Covariance: $\text{Cov}(X, Y) = \mathbb{E}[XY]$
   (c) Correlation: $\text{Corr}(X, Y) = \frac{\text{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$
6. Union bound: $\mathbb{P}(\bigcup_{i=1}^{n} X_i) \leq \sum_{i=1}^{n} \mathbb{P}(X_i)$
7. Iterated expectation/tower rule: $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$
8. Law of total variance: $\text{Var}(X) = \mathbb{E}[\text{Var}(X|Y)] + \text{Var}(\mathbb{E}[X|Y])$
9. Order statistics: the i-th largest value out of n realizations of X is distributed as $f_{X^{(i)}}(x) = n\binom{n-1}{i-1} f(x)F(x)^{i-1}(1 - F(x))^{n-i}$
10. Convolution: for a random variable $Z = X + Y$, $f_Z(z) = \int_{t=-\infty}^{\infty} f_X(z - t)f_Y(t)dt$

2 Problem Solving Techniques

1. It is important to be able to use counting as a means of generating probabilities: combinations, permutations, stars and bars (balls and bins), etc.
2. Be familiar with indicator variables and using them to calculate expectations and variances for various setups
3. Graphical density: reading a pdf from a graph and performing calculations with it
4. Probabilistic method: one way to show that something can be done is to show that simply randomly picking things is successful with nonzero probability

3 Moment Generating Functions

1. Moment generating functions are given by $M_x(s) = \mathbb{E}[e^{sX}]$; be able to recognize these for common distributions, and read off the parameters
2. The n-th moment of a random variable is $\mathbb{E}[X^n] = M_X^{(n)}(0) = \frac{d^n M_X}{ds^n}|_{s=0}$
3. For a linear combination $S = cA + dB$, the MGF of S is $M_S(s) = M_A(cs)M_B(ds)$, which is often simpler than computing the integral in a direct convolution (when A, B are independent)

4 Bounds/Concentration Inequalities

1. Markov's inequality: $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$ for nonnegative random variable X
2. Chebyshev's inequality: $\mathbb{P}(|X - \mathbb{E}[X]| \geq c) \leq \frac{\text{Var}(X)}{c^2}$
3. We can combine the (nonnegative) MGF with Markov's to get Chernoff's inequality: $\mathbb{P}(X \geq a) = \mathbb{P}(e^{sX} \geq e^{sa}) \leq \frac{\mathbb{E}[e^{sX}]}{[e^{sa}]}$, and taking the min over s to get the best bound

5 Convergence, Law of Large Numbers, CLT

1. Convergence in probability: $X_n \xrightarrow[n\to\infty]{\text{i.p.}} X$ if $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$, i.e. the probability that X_n deviates only from X goes to zero (but can still deviate infinitely)
   (a) Weak Law of Large Numbers (empirical mean converges to true mean in probability)
2. We can use CLT or bounds to set up confidence intervals for estimation

6 Applications

Note: auction theory can be found on past exams but is not in scope this semester.

1. Matrix sketching: instead of computing A^TB, compute the cheaper $(SA)^T(SB) = A^TS^TSB$ using a sketch matrix S, where $S^TS \approx I$
   (a) Understand expctation/variance calculations for elements of S^TS in the Gaussian-sketch and count-sketch setups
2. Fountain codes: sending messages robustly across a channel by XORing multiple packets together according to a distribution
   (a) Ideal soliton distribution: in expectation, always able to decode a packet at every step