In preparation for final

- Solutions to homeworks 7/8 will be posted today
- Sample final will be posted by tomorrow morning.
- I will have office hours Monday morning 8:30—10:30am.

Coverage–comprehensive, slight emphasis post-midterm

pre-mid: B&T ch 1-2

post-mid: B&T ch 3, 4.3, 5, 9.1, continuous, limits, mle.

everything in slides, hw

Mechanics

closed book, aside from one page of notes (8.5 x 11, both sides, handwritten)

I'm much more interested in setup and method than in numerical answers, so concentrate on giving a clear approach, perhaps including a terse english outline of your reasoning.

Corollary: calculators are probably irrelevant, but bring one to the exam if you want, just in case.

Format–similar to midterm:

T/F, multiple choice, problem-solving, explain, …

Story problems

counting principle (product rule)

permutations

combinations

indistinguishable objects

binomial coefficients

binomial theorem

partitions & multinomial coefficients

inclusion/exclusion

pigeon hole principle

chapter 1: axioms of probability

sample spaces & events

axioms

complements, Venn diagrams, deMorgan,
mutually exclusive events, etc.

equally likely outcomes

chapter 1: conditional probability and independence

conditional probability

chain rule, aka multiplication rule

total probability theorem

Bayes rule   yes, learn the formula

independence

conditional independence

gambler’s ruin

chapter 2: random variables

discrete random variables

probability mass function (pmf)

expectation of X

expectation of g(X) (i.e., a function of an r.v.)

linearity: expectation of X+Y and aX+b

variance

cumulative distribution function (cdf)

cdf as sum of pmf from -∞

independence; joint and marginal distributions

important examples: know pmf, mean, variance of these

bernoulli, binomial, poisson, geometric, uniform

some important (discrete) distributions

Name	PMF	$E[k]$	$E[k^2]$	σ^2
Uniform(a,b)	$f(k) = \frac{1}{(b-a+1)}, k = a, a+1, \ldots, b$	$\frac{a+b}{2}$		$\frac{(b-a+1)^2-1}{12}$
Bernoulli(p)	$f(k) = \begin{cases} 1-p & \text{if } k=0 \\ p & \text{if } k=1 \end{cases}$	p	p	$p(1-p)$
Binomial(p,n)	$f(k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n$	np		$np(1-p)$
Poisson(λ)	$f(k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, \ldots$	λ	$\lambda(\lambda+1)$	λ
Geometric(p)	$f(k) = p(1-p)^{k-1}, k = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{2-p}{p^2}$	$\frac{1-p}{p^2}$
Hypergeometric(n,N,m)	$f(k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, k = 0, 1, \ldots, N$	$\frac{nm}{N}$		$\frac{nm}{N}\left(\frac{(n-1)(m-1)}{N-1} + 1 - \frac{nm}{N}\right)$

See also the summary in B&T following pg 528

chapter 3: continuous random variables

especially 3.1–3.5

probability density function (pdf)

cdf as integral of pdf from $-\infty$

expectation and variance

Law of total probability, law of total expectation

Conditional expectation

Recall, in general, $E[f(X)] \neq f(E[X])$.

important examples

uniform, normal (incl Φ, "standardization"), exponential

know pdf and/or cdf, mean, variance of these

tail bounds

Markov

Chebyshev

Chernoff (lightly)

limit theorems

law of large numbers

central limit theorem

likelihood, parameter estimation, MLE (b&t 9.1)

likelihood

"likelihood" of observed data given a model

usually just a product of probabilities (or densities: "$\lim_{\delta \to 0} \ldots$"), by independence assumption

a function of (unknown?) parameters of the model

parameter estimation

if you know/assume the form of the model (e.g. normal, poisson,...), can you estimate the parameters based on observed data

many ways

maximum likelihood estimators

one way to do it–choose values of the parameters that maximize likelihood of observed data

method (usually) – solve
    "derivative (wrt parameter/s) of (log) likelihood = 0"

EM

iterative algorithm trying to find MLE in situations that are analytically intractable

usual framework: there are 0/1 hidden variables (e.g., from which component was this datum sampled) & problem much easier if they were known

E-step: given rough parameter estimates, find expected values of hidden variables

M-step: given expected values of hidden variables, find (updated) parameter estimates to maximize likelihood

Algorithm: iterate above alternately until convergence

probability & statistics, broadly

Noise, uncertainty & variability are pervasive

Learning to model it, derive knowledge, and compute despite it are critical

E.g., knowing the mean is valuable, but two scenarios with the same mean and different variances can behave very differently in practice.

want more?

Stat 390/1	probability & statistics
CSE 421	algorithms
CSE 427/8	computational biology
CSE 440/1	human/computer interaction
CSE 446	machine learning
CSE 472	computational linguistics
CSE 473	artificial intelligence

and others!

Thanks and Good Luck!