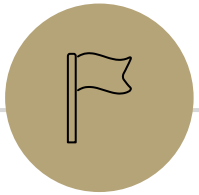
HW 7 solutions up front + in Allen 206

Victory Lap & Final Review

CSE 312 Summer 25
Lecture 23



Final Announcements

Final Logistics

Part 1: Thursday August 21st at 12pm in **DEM 104**

Part 2: Friday August 22nd at 12pm in **DEM 104**

Emergency comes up? Email Anna ASAP

Closed note, but reference sheet will be provided.

60 minutes per part

Full Details on website

https://courses.cs.washington.edu/courses/cse312/25su/exams/final.html

What will be on the final?

8 problems in total (2 of which are grading moral one for each part)

Multiple Choice

Probability

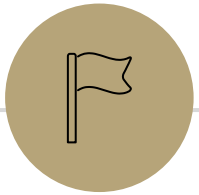
Zoo of Random Variables

Joint Distributions

Tail Bounds

Maximum Likelihood Estimation

$$\binom{6}{3} = 20$$



Victory Lap

What Have We Done?

Well let’s look back…

Content

Combinatorics (*fancy* counting)

Permutations, combinations, inclusion-exclusion, pigeonhole principle

Formal definitions for Probability

Probability space, events, conditional probability, independence, expectation, variance

Common patterns in probability

Equations and inequalities, “zoo” of common random variables, tail bounds

Continuous Probability

pdf, cdf, sample distributions, central limit theorem, estimating probabilities

Applications

Across CS, but with some focus on ML.

Themes

Precise mathematical communication

Both reading and writing dense statements.

Probability in the “real world”

A mix of CS applications

And some actual “real life” ones.

Refine your intuition

Most people have some base level feeling of what the chances of some event are.

We’re going to train you to have better gut feelings.

Use Your Powers Wisely

We’ve seen probability can be used in the real world!

But also that it:

Can be counter-intuitive/hard to explain (Bayes Rule/Real World)

Probability estimates can depend on the model you’re using (Real World)

Can be used to analyze ML applications, and think about the impacts of using them.

How (not to) lie with statistics

You now know a lot of the tools that people use to lie with statistics. (See also: INFO 270)

Some patterns to watch out for:

My smoke alarm is going off, please pay for my new house! (analogy from Matt Parker)

Make a model, find that an event that occurred had small probability/fails some statistical test, claim that the **only** explanation is something nefarious occurred.

Better response: could the model be wrong? Is this statistical test appropriate? Once in 100 year events do happen…about once in every hundred years, is this just the one?

How (not to) lie with statistics

See a story about testing?

Remember from Bayes' Rule that you need three numbers to understand a test. (3 of prior, posterior, false positive rate, false negative rate).

Headlines usually give you one number, that often isn't even one of the ones you need for Bayes ("this test is less accurate than a coin flip!").

The article itself, if you're lucky, might give you one or two of the numbers for Bayes – don't forget the prior!

How (not to) lie with statistics

We can apply our knowledge to the real world!

But if you're applying in a new domain, get information from domain experts, don't instantly assume because you know Bayes' Rule that you know better than domain experts.

Don't hesitate to use these tools to understand new domains better!

But do keep in mind some things can't be quantified and just because we can use an algorithm doesn't mean we always should.

What to take next?

ML (CSE 446) / NLP (CSE 447) using probability, linear algebra, and other techniques to extract patterns from data and make predictions.

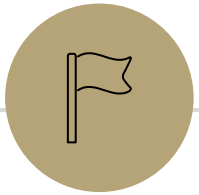
CSE 421 designing algorithms – very little direct probability, but the combinatorics we did at the beginning will be useful.

We also have a graduate level course in randomized algorithms, but it has a few more prereqs

CSE 426 Cryptography

CSE 422 Modern Algorithms

Other things!



Thank you!