

Victory Lap

CSE 312 23Su
Lecture 24

Announcement

Please fill out the course evaluation!

Show up to the exam on Friday 😊 The staff is available in office hours and over Ed to help you out until then.

My OH today are cancelled.

What's Next?

CSE 421: Algorithms (design algos – uses some combinatorics)

CSE 422: Toolkit for Modern Algorithms

CSE 426: Cryptography

CSE 427: Computational Biology

CSE 446: Machine Learning (probability + linear algebra)

CSE 447: Natural Language Processing

CSE 473: Artificial Intelligence (Bayes nets and such)

CSE 490Q: Quantum Computing

Once upon a time, there was: counting

Sum rule – choose from mutually exclusive (disjoint) sets

Product rule – decisions in a sequential process

Once upon a time, there was: counting

Complementary counting – sometimes, the complement is easier to count ($|A| = |\mathcal{U}| - |\bar{A}|$)

Inclusion-exclusion: choose from possibly overlapping sets

Once upon a time, there was: counting

Permutations – order distinct elements

***k*-permutation**

The number of *k*-element sequences of distinct symbols from a universe of *n* symbols is:

$$P(n, k) = n \cdot (n - 1) \cdots (n - k + 1) = \frac{n!}{(n - k)!}$$

Combinations – number of subsets of a set

***k*-combination**

The number of *k*-element subsets from a set of *n* symbols is:

$$C(n, k) = \frac{P(n, k)}{k!} = \frac{n!}{k! (n - k)!}$$

Once upon a time, there was: counting

Combinatorial proofs

Pigeonhole principle

If you have n pigeons and k pigeonholes, then there is **at least one** pigeonhole that has at least $\left\lceil \frac{n}{k} \right\rceil$ pigeons.

Stars-and-bars – indistinguishable objects into distinguishable bins

Counting sets leads to... discrete probability

Begin quantifying our uncertainty.

Probability Space

A (discrete) probability space is a pair $(\Omega, \mathbb{P})$ where:

Ω is the sample space

$\mathbb{P}: \Omega \rightarrow [0, 1]$ is the probability measure.

$\mathbb{P}$ satisfies:

- $\mathbb{P}(x) \geq 0$ for all x (*non-negativity*)
- $\sum_{x \in \Omega} \mathbb{P}(x) = 1$ (*normalization*)
- If $E, F \subseteq \Omega$ and $E \cap F = \emptyset$ then $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$ (*countable additivity*)

Counting sets leads to... discrete probability

Equally likely outcomes

Corollaries: complementation, monotonicity, inclusion-exclusion

- $\mathbb{P}(\bar{E}) = 1 - \mathbb{P}(E)$ (complementation)
- If $E \subseteq F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$ (monotonicity)
- $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$ (inclusion-exclusion)

Sometimes we have extra information...

Restrict the sample space

Bayes uses a prior to calculate a posterior

Sometimes we have extra information...

Law of Total Probability

Chain rule

Chain Rule

$$\begin{aligned} & \mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \\ &= \mathbb{P}(A_n | A_1 \cap \cdots \cap A_{n-1}) \cdot \mathbb{P}(A_{n-1} | A_1 \cap \cdots \cap A_{n-2}) \cdots \mathbb{P}(A_2 | A_1) \cdot \mathbb{P}(A_1) \end{aligned}$$

Sometimes that extra is irrelevant...

Independence

Pairwise Independence

Events $A_1, A_2, \dots, A_n$ are pairwise independent if
$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j) \text{ for all } i, j \text{ where } i \neq j$$

Mutual Independence

Events $A_1, A_2, \dots, A_n$ are mutually independent if
$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k})$$

for every subset $\{i_1, i_2, \dots, i_k\}$ of $\{1, 2, \dots, n\}$.

Sometimes that extra is irrelevant...

Conditional independence

We say A and B are conditionally independent on C if

$$\mathbb{P}(A \cap B|C) = \mathbb{P}(A|C) \cdot \mathbb{P}(B|C)$$

Summarize Numerical Info from Ω

(Discrete) Random variables

- > Range
- > PMF
- > CDF
- > Expectation

Summarize Numerical Info from Ω

(Discrete) Random variables

> Linearity of Expectation

Summarize Numerical Info from Ω

(Discrete) Random variables

> Variance

Variance

The variance of a random variable X is

$$\text{Var}(X) = \sum_{\omega} \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Summarize Numerical Info from Ω

$X \sim \text{Unif}(a, b)$

$$p_X(k) = \frac{1}{b - a + 1}$$

$$\mathbb{E}[X] = \frac{a + b}{2}$$

$$\text{Var}(X) = \frac{(b - a)(b - a + 2)}{12}$$

$X \sim \text{Ber}(p)$

$$p_X(0) = 1 - p;$$

$$p_X(1) = p$$

$$\mathbb{E}[X] = p$$

$$\text{Var}(X) = p(1 - p)$$

$X \sim \text{Bin}(n, p)$

$$p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$\mathbb{E}[X] = np$$

$$\text{Var}(X) = np(1 - p)$$

$X \sim \text{Geo}(p)$

$$p_X(k) = (1 - p)^{k-1} p$$

$$\mathbb{E}[X] = \frac{1}{p}$$

$$\text{Var}(X) = \frac{1 - p}{p^2}$$

$X \sim \text{NegBin}(r, p)$

$$p_X(k) = \binom{k-1}{r-1} p^r (1 - p)^{k-r}$$

$$\mathbb{E}[X] = \frac{r}{p}$$

$$\text{Var}(X) = \frac{r(1 - p)}{p^2}$$

$X \sim \text{HypGeo}(N, K, n)$

$$p_X(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

$$\mathbb{E}[X] = n \frac{K}{N}$$

$$\text{Var}(X) = \frac{K(N - K)(N - n)}{N^2(N - 1)}$$

$X \sim \text{Poi}(\lambda)$

$$p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\mathbb{E}[X] = \lambda$$

$$\text{Var}(X) = \lambda$$

Our sample spaces may be uncountable!

Continuous RVs

PDFs

$$F_X(k) = \mathbb{P}(X \leq k) = \int_{-\infty}^k f_X(z) \, dz$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(z) \cdot f_X(z) \, dz$$

Our sample spaces may be uncountable!

$$X \sim \text{Unif}(a, b)$$

$$f_X(k) = \frac{1}{b-a}$$
$$\mathbb{E}[X] = \frac{a+b}{2}$$
$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$X \sim \text{Exp}(\lambda)$$

$$f_X(k) = \lambda e^{-\lambda k} \text{ for } k \geq 0$$
$$\mathbb{E}[X] = \frac{1}{\lambda}$$
$$\text{Var}(X) = \frac{1}{\lambda^2}$$

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

$$f_X(k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$
$$\mathbb{E}[X] = \mu$$
$$\text{Var}(X) = \sigma^2$$

An Approximation...

Central Limit Theorem

Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables, with mean μ and variance σ^2 . Let $Y_n = \frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$

As $n \rightarrow \infty$, the CDF of Y_n converges to the CDF of $\mathcal{N}(0, 1)$

CLT E.g.

More than 1 RV? You got it!

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x, y) = P(X = x, Y = y)$	$f_{X,Y}(x, y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x, y) = \sum_{t \leq x} \sum_{s \leq y} p_{X,Y}(t, s)$	$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$
Normalization	$\sum_x \sum_y p_{X,Y}(x, y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x, y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
Expectation	$E[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y)$	$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$
Conditional PMF/PDF	$p_{X Y}(x y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$	$f_{X Y}(x y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$
Conditional Expectation	$E[X Y = y] = \sum_x x p_{X Y}(x y)$	$E[X Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x y) dx$
Independence	$\forall x, y, p_{X,Y}(x, y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x, y) = f_X(x) f_Y(y)$

More than 1 RV? You got it!

Calculating probabilities

Law of Total Expectation

Let $A_1, A_2, \dots, A_k$ be a partition of the sample space, then

$$\mathbb{E}[X] = \sum_{i=1}^k \mathbb{E}[X|A_i] \mathbb{P}(A_i)$$

Variance of the sum of RVs?

Covariance

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\text{Var}(X + Y) = \text{Var}(X) + 2 \cdot \text{Cov}(X, Y) + \text{Var}(Y)$$

What if we can't/don't need an exact probability calculation?

Tail bounds / concentration inequalities

Markov's Inequality

Let X be a random variable supported (only) on non-negative numbers. For any $t > 0$

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$$

Chebyshev's Inequality

Let X be a random variable. For any $t > 0$

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\text{Var}(X)}{t^2}$$

What if we can't/don't need an exact probability calculation?

Tail bounds / concentration inequalities

(Multiplicative) Chernoff Bound

Let $X_1, X_2, \dots, X_n$ be *independent* Bernoulli random variables.

Let $X = \sum X_i$, and $\mu = \mathbb{E}[X]$. For any $0 \leq \delta \leq 1$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq \exp\left(-\frac{\delta^2\mu}{3}\right) \text{ and } \mathbb{P}(X \leq (1 - \delta)\mu) \leq \exp\left(-\frac{\delta^2\mu}{2}\right)$$

What if we can't/don't need an exact probability calculation?

Union Bound

For any events E, F
$$\mathbb{P}(E \cup F) \leq \mathbb{P}(E) + \mathbb{P}(F)$$

What if we have data but not the model parameters?

Maximum Likelihood Estimation

Likelihood

The likelihood of i.i.d observations $x_1, \dots, x_n$ from a discrete distribution is

$$\mathcal{L}(x_1, \dots, x_n ; \theta) = \prod_{i=1}^n \mathbb{P}(x_i; \theta)$$

What if we have data but not the model parameters?

Maximum Likelihood Estimation

An estimator $\hat{\theta}$ is “unbiased” if

$$\mathbb{E}[\hat{\theta}] = \theta$$

An estimator $\hat{\theta}$ is “consistent” if

$$\lim_{n \rightarrow \infty} \mathbb{E}[\hat{\theta}] = \theta$$

Some Extras! 😊

Naïve Bayes Spam Filtering

Bloom Filters

Polling

Algorithmic Privacy

Markov Chains (and Markov Chain Monte Carlo algorithms)

Multi-armed Bandits

Python + Latex!

Your Takeaways 😊

Better precise mathematical communication.

Strong problem-solving skills.

A solid foundation to learn more about algorithms, machine learning, data structures, networks, and so on!

Your Takeaways 😊

Better intuition (but more often, knowledge that intuition often fails in the probabilistic world)

Know that people often use statistics as a tool to support their (dubious) claims. There are a lot of different tricks, but here's my one-stop method: if you don't completely understand what a number is and where it came from, then you can't be sure of what it means.

Just because something **can** be quantified doesn't mean we **should**.

Thank you for a wonderful quarter!