Review.

After Midterm content.

See lecture 13, for pre-Midterm review.

Counting.

First Rule: Enumerate objects with sequence of choices.
  Number of Objects: $n_1 \times n_2 \dots$.
  Example: Poker deals.

Second Rule: Divide out if by ordering of same objects.
  Example: Poker hands. Orderings of ANAGRAM.

Sum Rule: If sets of objects disjoint add sizes.
  Example: Hands with joker, hands without.

Inclusion/Exclusion: For arbtrary sets A, B.
  $|A \cup B| = |A| + |B| - |A \cap B|$
  Example: 10 digit numbers with 9 in the first or second digit.

Combinatorial Proofs.

Theorem: $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$.
Proof: How many size k subsets of $n+1$? $\binom{n+1}{k}$.

How many size k subsets of $n+1$?
How many contain the first element?
 Choose first element,
  need to choose $k-1$ more from remaining n elements.
$\implies \binom{n}{k-1}$

How many don't contain the first element ?
 Need to choose k elements from remaining n elts.
$\implies \binom{n}{k}$

So, $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$. $\square$

Uncountability/Undecidability.

Natural Numbers are countable. Definition.

Rationals are countable cuz bijection.

Reals are not.

Why? Diagonalization.

Halt is undecidable.

Why? Diagonalization.

Reductions **from** Halt give more undecidable problems.

 Reductions use program for problem A to solve HALT.

Concept: Can programatically modify *text* of input program (to HALT).
Concept: Can call program A.

CS70: Review of Probability.

Probability Review

Discrete Probability

Probability Space: Ω, $Pr[\omega] \geq 0$, $\sum_{\omega} Pr[\omega] = 1$.

Random Variable: Function on Sample Space: $X : \Omega \to R$.

Distribution: Function $Pr[X = a] \geq 0$. $\sum_a Pr[X = a] = 1$.

Expectation: $E[X] = \sum_{\omega} X(\omega) \times Pr[\omega] = \sum_a a \times Pr[X = a]$.

Many Random Variables: each one function on a sample space.

Joint Distributions: Function $Pr[X = a, Y = b] \geq 0$.
$\sum_{a,b} Pr[X = a, Y = b] = 1$.

Linearity of Expectation: $E[X + Y] = E[X] + E[Y]$.

Applications: compute expectations by decomposing.

  Indicators: Empty bins, Fixed points.
  Time to Coupon: Sum times to "next" coupon.
   Geometric distribution vs. direct.
  Birthday Paradox with expectation (and without.)

$Y = f(X)$ is Random Variable.
 Distribution of Y from distribution of X:
$Pr[Y = y] = \sum_{x \in f^{-1}y} Pr[X = x]$.

Tail Bounds: WWLN

$\boxed{\text{Variance}}$

- Variance: $var[X] := E[(X-E[X])^2] = E[X^2] - E[X]^2$
- Fact: $var[aX+b] = a^2 var[X]$
- Sum: X,Y,Z pairwise ind. $\Rightarrow var[X+Y+Z] = \cdots$
- Markov: $Pr[X \geq a] \leq E[f(X)]/f(a)$ where ...
- Chebyshev: $Pr[|X-E[X]| \geq a] \leq var[X]/a^2$
- WLLN: X_m i.i.d. $\Rightarrow \frac{X_1+\cdots+X_n}{n} \approx E[X]$

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega\in\Omega} Pr(w) = 1$.
Random Variables: $X : \Omega \to R$.
Associated event: $Pr[X=a] = \sum_{\omega:X(\omega)=a} Pr(\omega)$
X and Y independent $\iff$ all associated events are independent.
Expectation: $E[X] = \sum_a aPr[X=a] = \sum_{\omega\in\Omega} X(\omega)Pr(\omega)$.
Linearity: $E[X+Y] = E[X]+E[Y]$.

Variance: $Var(X) = E[(X-E[X])^2] = E[X^2]-(E(X))^2$
For independent X,Y, $Var(X+Y) = Var(X)+Var(Y)$.
Also: $Var(cX) = c^2 Var(X)$ and $Var(X+b) = Var(X)$.

Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $Var(X) = \lambda$.
Binomial: $X \sim B(n,p)$ $E(X) = np$, $Var(X) = np(1-p)$
Uniform: $X \sim U\{1,\ldots,n\}$ $E[X] = \frac{n+1}{2}$, $Var(X) = \frac{n^2-1}{12}$.
Geometric: $X \sim G(p)$ $E(X) = \frac{1}{p}$, $Var(X) = \frac{1-p}{p^2}$

Continuous Probability

1. pdf: $Pr[X \in (x, x+\delta]] = f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = \int_{-\infty}^{x} f_X(y)dy$.
3. $U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1-\exp\{-\lambda x\}$ for $x \geq 0$.
5. Target: $f_X(x) = 2x1\{0 \leq x \leq 1\}; F_X(x) = x^2$ for $0 \leq x \leq 1$.
6. Joint pdf: $Pr[X \in (x, x+\delta), Y = (y, y+\delta)) = f_{X,Y}(x,y)\delta^2$.
   6.1 Conditional Distribution: $f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.
   6.2 Independence: $f_{X|Y}(x,y) = f_X(x)$

Continuous Probability: Moments.

- Expectation.
  $E[X] = \int_{-\infty}^{\infty} xf(x)dx$.
- Variance.
  $E[(X-E[X])^2]$ Same as discrete.
- Markov? Sure.
- Chebshev? Sure.

Distributions.

- $X \sim U[a,b]$
  $f_X(x) = \frac{1}{(b-a)} 1\{x \in [a,b]\}$. $F(x) = \min(\frac{x-a}{b-a} 1\{x \in [a,b]\}, 1.0)$
  $E[X] = \frac{b-a}{2}$.
  $Var(X) = \frac{(b-a)^2}{12}$.
- $X \sim Expo(\lambda)$
  $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$ $F_X(x) = 1 - e^{\lambda x}$.
  $E[X] = \frac{1}{\lambda}$.
  $Var[X] = \frac{1}{\lambda^2}$
  $X = lim_{n\to\infty} \frac{1}{n} X\ G(\lambda/n)$
  $Pr[X > s+t | X > s] = Pr[X > t]$: Memoryless.
- $X \sim N(\mu,\sigma)$. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2}2$ $F(x) = \Phi(x) = \int_{-\infty}^{x} f(x)dx$
  CLT: $A_n = \frac{X_1+\cdots X_n}{n}$, $E[X] = \mu$, and $Var(X) = \sigma^2$.
  $lim_{n\to\infty} A_n \to N(\mu,\sigma)$
  $\frac{X-\mu}{\sigma} \sim N(0,1)$.

Markov Chains

- Markov Chain: $Pr[X_{n+1} = j | X_0, \ldots, X_n = i] = P(i,j)$
- FSE: $\beta(i) = 1 + \sum_j P(i,j)\beta(j); \alpha(i) = \sum_j P(i,j)\alpha(j)$.
- $\pi_n = \pi_0 P^n$
- π is invariant iff $\pi P = \pi$
- Irreducible $\Rightarrow$ one and only one invariant distribution π
- Irreducible $\Rightarrow$ fraction of time in state i approaches $\pi(i)$
- Irreducible + Aperiodic $\Rightarrow \pi_n \to \pi$.
- Calculating π: One finds $P\pi = \pi$ and π is distribution.

Some exercise.

1. True or False
2. Some Key Results
3. Quiz 1: G
4. Quiz 2: PG
5. Quiz 3: R
6. Common Mistakes

True or False

- Ω and A are independent. True
- $Pr[A \cap B] = Pr[A] + Pr[B] - Pr[A \cup B]$. True
- $Pr[A \setminus B] \geq Pr[A] - Pr[B]$. True
- $X_1, \ldots, X_n$ i.i.d. $\implies var(\frac{X_1 + \cdots + X_n}{n}) = var(X_1)$. False: $\times \frac{1}{n}$
- $Pr[|X - a| \geq b] \leq \frac{E[(X-a)^2]}{b^2}$. True
- $X_1, \ldots, X_n$ i.i.d. $\implies \frac{X_1 + \cdots + X_n - nE[X_1]}{n\sigma(X_1)} \to \mathcal{N}(0,1)$. False: $\sqrt{n}$
- $X = Expo(\lambda) \implies Pr[X > 5 | X > 3] = Pr[X > 2]$. True:
  $\frac{\exp\{-\lambda 5\}}{\exp\{-\lambda 3\}} = \exp\{-\lambda 2\}$.

Correct or not?

$A_n = \frac{\sum_{i=0}^{n} X_i}{n}$, X_i i.i.d, mean μ and variance σ.
When $n \gg 1$, one has

- $A_n \in [\mu - 2\sigma\frac{1}{n}, \mu + 2\sigma\frac{1}{n}]$ with prob $\geq 95\%$. No
- $A_n \in [\mu - 2\sigma\frac{1}{\sqrt{n}}, \mu + 2\sigma\frac{1}{\sqrt{n}}]$ with prob $\geq 95\%$. Yes
- If $0.3 < \sigma < 3$, then
  $A_n \in [\mu - 0.6\frac{1}{\sqrt{n}}, \mu + 0.6\frac{1}{\sqrt{n}}]$ with prob $\geq 95\%$. No
- If $0.3 < \sigma < 3$, then
  $A_n \in [\mu - 6\frac{1}{\sqrt{n}}, \mu + 6\frac{1}{\sqrt{n}}]$ with prob $\geq 95\%$. Yes

Match Items

[1] $Pr[X \geq z] \leq \frac{E[f(X)]}{f(a)}$.

[2] $Pr[|X - E[X]| \geq z] \leq \frac{Var(X)}{z^2}$.

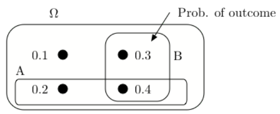
[3] $\sum_y yPr[Y = y | X = x]$.

[4] $Pr[|\frac{X_1 + X_2 + \cdots + X_n}{n} - E[X_1]| \geq \varepsilon] \to 0$

[5] $Pr[|\frac{X_1 + X_2 + \cdots + X_n}{n} - E[X_1]| \leq \varepsilon] \to 1$

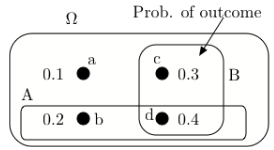
- WLLN (4) and (5)
- Chebyshev (2)
- Markov's inequality (1)
- $E[Y|X = x]$ (3)

Quiz 1: G



1. What is $P[A|B]$?
   $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{0.4}{0.7}$
2. What is $Pr[B|A]$?
   $Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]} = \frac{0.4}{0.6}$
3. Are A and B positively correlated?
   No. $Pr[A \cap B] = 0.4 < Pr[A]Pr[B] = 0.6 \times 0.7$.

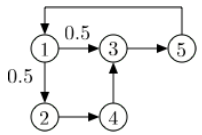
Quiz 1: G



ω	$X(\omega)$	$Y(\omega)$
a	0	0
b	1	0
c	0	2
d	1	2

4. What is $cov(X, Y)$?
   $cov(X, Y) = E[XY] - E[X]E[Y]$
   $= 0.8 - 0.6 \times 1.4 = -0.04$

Quiz 1: G

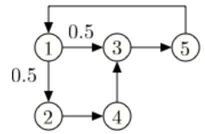


7. Is this Markov chains irreducible? Yes.
8. Is this Markov chain periodic?

   No. The return times to 3 are $\{3, 5, ..\}$: coprime!
9. Does π_n converge to a value independent of π_0? Yes!
10. Does $\frac{1}{n}\sum_{m=1}^{n-1} 1\{X_m = 3\}$ converge as $n \to \infty$? Yes!
11. Calculate π.

    Let $a = \pi(1)$. Then $a = \pi(5), \pi(2) = 0.5a$,
    $\pi(4) = \pi(2) = 0.5a, \pi(3) = 0.5\pi(1) + \pi(4) = a$.
    Thus, $\pi = [a, 0.5a, a, 0.5a, a] = [1, 0.5, 1, 0.5, 1]a$, so $a = 1/4$.

Quiz 1: G



12. Write the first step equations for calculating the mean time from 1 to 4.

$$\beta(1) = 1 + 0.5\beta(2) + 0.5\beta(3)$$
$$\beta(2) = 1$$
$$\beta(3) = 1 + \beta(5)$$
$$\beta(5) = 1 + \beta(1).$$

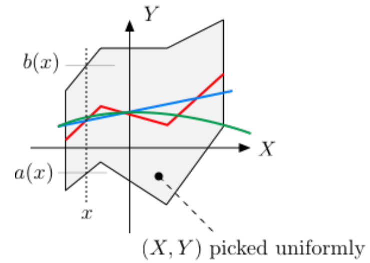
13. Solve these equations.

$$\begin{aligned}\beta(1) &= 1 + 0.5 \times 1 + 0.5 \times (1 + (1 + \beta(1))) \\ &= 2.5 + 0.5\beta(1).\end{aligned}$$

Hence, $\beta(1) = 5$.

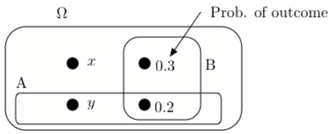
Quiz 1: G

14. Which is $E[Y|X]$? Blue, red or green?



Answer: Red.
Given $X = x$, $Y = U[a(x), b(x)]$. Thus, $E[Y|X = x] = \frac{a(x)+b(x)}{2}$.

Quiz 2: PG



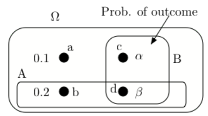
1. Find (x, y) so that A and B are independent.

   We need
$$Pr[A \cap B] = Pr[A]Pr[B]$$
   That is,
$$0.2 = (y + 0.2) \times 0.5 = 0.5y + 0.1$$
   Hence,
$$y = 0.2 \text{ and } x = 0.3.$$

2. Find the value of x that maximizes $Pr[B|A]$.

   When $x = 0.5$, $Pr[B|A] = 1$.

Quiz 2: PG



ω	$X(\omega)$	$Y(\omega)$
a	0	0
b	1	0
c	0	2
d	1	2

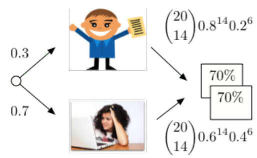
3. Find α so that X and Y are independent.

   We need
$$Pr[X = 0, Y = 0] = Pr[X = 0]Pr[Y = 0]$$
   That is,
$$0.1 = (0.1 + \alpha) \times (0.1 + 0.2) = 0.03 + 0.3\alpha$$
   Hence,
$$\alpha = 0.233$$
   Typically: check $Pr[X = 2, Y = 0] = Pr[X = 2]Pr[Y = 0]$ and so on.

   But: A and B independent $\iff \bar{A}, B$ independent.
   Take: $A =$ "$X = 0$" and $B =$ "$Y = 0$", since only two values for X, Y

Quiz 2: PG

4. A CS70 student is great w.p. 0.3 and good w.p. 0.7. A great student solves each question correctly w.p. 0.8 whereas a good student does it w.p. 0.6. One student got right 70% of the 10 questions on Midterm 1 and 70% of the 10 questions on Midterm 2. What is the expected score of the student on the final?



$$p := Pr[\text{great}|\text{scores}] = \frac{0.3\binom{20}{14}0.8^{14}0.2^{6}}{0.3\binom{20}{14}0.8^{14}0.2^{6} + 0.7\binom{20}{14}0.6^{14}0.4^{6}}$$
$$= \frac{(0.3)0.8^{14}0.2^{6}}{(0.3)0.8^{14}0.2^{6} + (0.7)0.6^{14}0.4^{6}} \approx 0.27$$

Expected score $= p80\% + (1-p)60\% \approx 65\%$.

Quiz 2: PG

6. You roll a balanced six-sided die 20 times. Use CLT to upper-bound the probability that the total number of dots exceeds 85.

Let $X = X_1 + \cdots + X_{20}$ be the total number of dots.
Then
$$\frac{X-70}{\sigma\sqrt{20}} \approx \mathcal{N}(0,1)$$
where
$$\sigma^2 = var(X_1) = (1/6)\sum_{m=1}^{6} m^2 - (3.5)^2 \approx 2.9 = 1.7^2.$$
Now,
$$\begin{aligned} Pr[X > 85] &= Pr[X - 70 > 15] \\ &= Pr[\frac{X-70}{1.7 \times 4.5} > \frac{15}{1.7 \times 4.5}] \\ &= Pr[\frac{X-70}{1.7 \times 4.5} > 2] \approx 2.5\%. \end{aligned}$$

Quiz 2: PG

7. You roll a balanced six-sided die 20 times. Use Chebyshev to upper-bound the probability that the total number of dots exceeds 85.

Let $X = X_1 + \cdots + X_{20}$ be the total number of dots.
Then
$$\begin{aligned} Pr[X > 85] &= Pr[X - 70 > 15] \le Pr[|X - 70| > 15] \\ &\le \frac{var(X)}{15^2}. \end{aligned}$$
Now,
$$var(X) = 20 var(X_1) = 20 \times 2.9 = 58.$$
Hence,
$$Pr[X > 85] \le \frac{58}{15^2} \approx 0.26.$$

Quiz 3: R

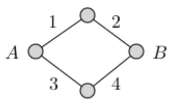
1. The lifespans of good lightbulbs are exponentially distributed with mean 1 year. Those of defective bulbs are exponentially distributed with mean 0.8. All the bulbs in one batch are equally likely to be good or defective. You test one bulb and note that it burns out after 0.6 year. (a) What is the probability you got a batch of good bulbs? (b) What is the expected lifespan of another bulb in that batch?

   Hint: If $X = Expo(\lambda), f_X(x) = \lambda e^{-\lambda x} 1\{x > 0\}, E[X] = 1/\lambda$.

Let X be the lifespan of a bulb, G the event that it is good, and B the event that it is bad.

$$\begin{aligned} (a) \quad & p := Pr[G|X \in (0.6, 0.6+\delta)] \\ &= \frac{0.5 Pr[X \in (0.6, 0.6+\delta)|G]}{0.5 Pr[X \in (0.6, 0.6+\delta)|G] + 0.5 Pr[X \in (0.6, 0.6+\delta)|D]} \\ &= \frac{e^{-0.6}\delta}{e^{-0.6}\delta + (0.8)^{-1} e^{-(0.8)^{-1} 0.6}\delta} \approx 0.488. \\ (b) \quad & E[\text{ lifespan of other bulb }] = p \times 1 + (1-p) \times 0.8 \approx 0.9. \end{aligned}$$

Quiz 3: R



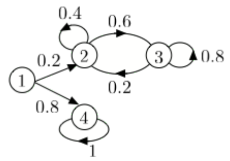
2. In the figure, $1, 2, 3, 4$ are links that fail after i.i.d. times that are $U[0,1]$.

   Find the average time until A and B are disconnected.

Let X_k be the lifespan of link k, for $k = 1, \ldots, 4$.
We are looking for $E[Z]$ where $Z = \max\{Y_1, Y_2\}$ with $Y_1 = \min\{X_1, X_2\}$ and $Y_2 = \min\{X_3, X_4\}$.

$$\begin{aligned} Pr[Y_1 > t] &= Pr[X_1 > t] Pr[X_2 > t] = (1-t)^2 \\ Pr[Z \le t] &= Pr[Y_1 \le t] Pr[Y_2 \le t] = (1 - (1-t)^2)^2 \\ &= (2t - t^2)^2 = 4t^2 - 4t^3 + t^4 \\ f_Z(t) &= 8t - 12t^2 + 4t^3 \\ E[Z] &= \int_0^1 t f_Z(t) dt = 8\frac{1}{3} - 12\frac{1}{4} + 4\frac{1}{5} \\ &\approx 0.4667. \end{aligned}$$

Quiz 3: R



3. We are given π_0. Find $\lim_{n\to\infty} \pi_n$.
   With probability $\alpha := 0.2\pi_0(1) + \pi_0(2) + \pi_0(3)$, the MC ends up in $\{2,3\}$.
   With probability $1 - \alpha$, it ends up in state 4.
   If it is in $\{2,3\}$, the probability that it is in state 2 converges to
   $$\frac{0.2}{0.2 + 0.6} = 0.25.$$
   Hence, the limiting distribution is
   $$[0, 0.25\alpha, 0.75\alpha, 1-\alpha].$$

Quiz 3: R

4. A bag has n red and n blue balls. You pick two balls (no replacement). Let $X = 1$ if ball 1 is red and $X = -1$ otherwise. Define Y likewise for ball 2.
   $\rightarrow$ Are X and Y positively, negatively, or un- correlated?
   Clearly, negatively.
5. Calculate $cov(X,Y)$.
   $cov(X,Y) = E[XY] - E[X]E[Y]$
   $E[X] = E[Y]$, by symmetry
   $E[X] = 0$
   $E[XY] = Pr[X = Y] - Pr[X \neq Y] = 2Pr[X = Y] - 1$
   $Pr[X = Y] = (n-1)/(2n-1)$
   E.g., if $X = +1$ = red, then Y is red w.p. $(n-1)/(2n-1)$
   $E[XY] = 2(n-1)/(2n-1) - 1 = -1/(2n-1) = cov(X,Y)$.

Common Mistakes

- $\Omega = \{1,2,3\}$. Define X, Y with $cov(X,Y) = 0$ and X, Y not independent.

  Let $X = 0, Y = 1$. No: They are independent.

  Let $X(1) = -1, X(2) = 0, X(1) = 1, Y(1) = 0, Y(2) = 1, Y(3) = 0$.
- $3 \times 3.5 = 12.5$. No.
- $E[X^2] = E[X]^2$. No.
- $X = B(n,p) \implies var(X) = n^2 p(1-p)$. No.
- $Pr[X = a] = Pr[X = a|A] + Pr[X = a|\bar{A}]$. No.
- $\sum_{n=0}^{\infty} a^n = 1/a$. No.
- CS70 is difficult. No.
- I will do poorly on the final. No.

Thanks and Best Wishes!