CSM Final Review

:)

Discrete Math

Fahad, Katya

Propositions

1. [True/False]: $\forall x\,(P(x) \wedge Q(x))$ is equivalent to $(\forall x, P(x)) \wedge (\forall x, Q(x))$. Explain.
2. [True/False]: $(\forall n \in N)(P(n) \Rightarrow Q(n)) \Rightarrow (\neg \exists n \in N)(Q(n) \Rightarrow \neg P(n))$ for all P,Q? Explain.
3. Suppose $P(k) \Rightarrow P(k+2)$. For the following, label as always true, sometimes true, never true:
   a. $P(0) \Rightarrow \forall n\, P(n+1)$
   b. $\exists n\ \exists m>n$ s.t $[P(2n) \wedge \neg P(2m)]$

Proposition Solutions

1. True or False: $\forall x\,(P(x) \wedge Q(x))$ is equivalent to $(\forall x, P(x)) \wedge (\forall x, Q(x))$. Explain.
   True. Recall that to prove $A \Leftrightarrow B$, we have to prove both $A \Rightarrow B$ and $B \Rightarrow A$.

   Suppose the first statement $\forall x\,(P(x) \wedge Q(x))$ is true. This means for all x, $P(x) \wedge Q(x)$ is true, so $P(x)$

   and $Q(x)$ are both true. Thus, the statement $(\forall x, P(x))$ is true, and similarly the statement $(\forall x, Q(x))$ is true, so the conjunction $(\forall x, P(x)) \wedge (\forall x, Q(x))$ is true.

   Conversely, suppose the second statement $(\forall x, P(x)) \wedge (\forall x, Q(x))$ is true. This means for all x, both $P(x)$ is true and $Q(x)$ is true, which implies $P(x) \wedge Q(x)$ is true. Thus, the first statement $\forall x\,(P(x) \wedge Q(x))$ is true. This completes the proof of the equivalence

2. $(\forall n \in N)(P(n) \Rightarrow Q(n)) \Rightarrow (\neg \exists n \in N)(Q(n) \Rightarrow \neg P(n))$:
   False, assume both $P(n)$ and $Q(n)$ are are false for all n

Propositions Solutions (contd)

3. Suppose $P(k) \Rightarrow P(k+2)$. For the following, label as always true, sometimes true, never true:
   a. $P(0) \Rightarrow \forall n\, P(n+1)$

      Sometimes. P(2) doesn't need to imply P(3)

   b. $\exists n \exists m > n$ s.t $[P(2n) \wedge \neg P(2m)]$

      False, by the assumption, if one even number is true, then every other even number greater is true as well.

Stable Marriage

[True/False]: If w is the top choice of m and m is the top choice of w, then m and w must be paired with each other in any stable matching

[True/False]: Suppose that in an instance of the original Stable Marriage problem with n couples (so that every man ranks every woman and vice versa), there is a man M who is last on each woman's list and a woman W who is last on every man's list. If the Gale-Shapley algorithm is run on this instance, than M and W will be paired with each other.

[True/False]: Suppose we have an instance of the original Stable Marriage problem with M and W as above. In *any* stable solution to the instance, M and W will be paired with each other.

[True/False]: Consider an alternative to the Propose and Reject algorithm (with no rejections), where women take turns choosing the best available husband from the remaining unchosen men. On day 1, the oldest woman chooses her most preferred man, and marries him. On day k, the k-th eldest woman chooses her most preferred choice from the remaining unmarried men, and marries him. No matter what the preferences are, this process always results in a stable matching.

Given m men and m women, for any $m \geq 2$, what is the minimum number of stable pairings that must exist for any sets of preferences? Justify your answer with a specific example attaining the minimum

Stable Marriage Solutions

[**True**/False]: If w is the top choice of m and m is the top choice of w, then m and w must be paired with each other in any stable matching

[**True**/False]: Suppose that in an instance of the original Stable Marriage problem with n couples (so that every man ranks every woman and vice versa), there is a man M who is last on each woman's list and a woman W who is last on every man's list. If the Gale-Shapley algorithm is run on this instance, than M and W will be paired with each other.
To prove it directly, note that no man will propose to W unless he has been rejected by all n-1 other women. Let X be the first man to propose to W. At the time this happens, all the other women must be engaged to men they prefer to X. This can only happen if X is M, as otherwise M would have been preferred to X by some woman. So M proposes to W, she accepts and then everyone is engaged and the algorithm stops.

[**True**/False]: Suppose we have an instance of the original Stable Marriage problem with M and W as above. In *any* stable solution to the instance, M and W will be paired with each other.
If M is instead married to W' and W is married to M', then M' and W' each prefer each other to their assigned spouse, and the matching is not stable. So if M and W are not married to each other, the matching is not stable.

[True/**False**]: Consider an alternative to the Propose and Reject algorithm (with no rejections), where women take turns choosing the best available husband from the remaining unchosen men. On day 1, the oldest woman chooses her most preferred man, and marries him. On day k, the k-th eldest woman chooses her most preferred choice from the remaining unmarried men, and marries him. No matter what the preferences are, this process always results in a stable matching.
Let the M be denoted by 1, 2 and the women by A, B. Assume that their preference lists are as follows:

1: B, A	A: 1, 2
2: A, B	B: 1, 2

If we run the new SMA on this, we will get the following pairs: (A, 1), (B, 2). This pairing is unstable. We found a counterexample

Given m men and m women, for any m≥2, what is the minimum number of stable pairings that must exist for any sets of preferences? Justify your answer with a specific example attaining the minimum
One pairing: We know that the Stable Marriage algorithm always terminates with one stable pairing, so we need to show that there is a set of preferences for which there is only one stable pairing. Let the Men be $M_1,\ldots,M_m$ and the Women be $W_1,\ldots,W_m$. Suppose that for every $1 \le i \le m$, M_i's top-ranked person is W_i, and W_i's top-ranked person is M_i. Then the pairing that matches up all (M_i,W_i) is the onlystable pairing, because if M_i and W_i are not paired, they will be a rogue couple

Mod Arithmetic

1. Find the inverse of 7 mod 48.
2. Let p,q be distinct prime numbers. Prove that $p^{q-1} + q^{p-1} = 1 \pmod{pq}$

Mod Solutions

1. Find the inverse of 7 mod 48.
   Notice that $7 * 7 = 49 = 1 \bmod 48$, so the inverse of 7 is itself.

2. Prove $p^{q-1} + q^{p-1} = 1 \pmod{pq}$
   Since p,q are distinct prime numbers, their gcd is 1. So, we can apply Fermat's Little Theorem a few times. Firstly, $p^{(q-1)}$ is 1 mod q, while $q^{(p-1)}$ is 0 mod q. So, their sum is equal to 1 mod p. Similarly, $q^{(p-1)}$ mod p is 1, while $p^{(q-1)}$ is 0 mod p, and hence, their sum mod p is 1 as well. So, in the end, we get that $q^{(p-1)} + p^{q-1} - 1$ is divisible by both p and by q, and since the two numbers are prime, they are divisible by pq. So the above equation is equal to 0 mod pq, and adding 1 to each side gets us what we need.

RSA

Fahad's Fault

Suppose, CSM is using RSA with modulus *n* and public exponent *e*. One day they are hacked, and their private key *d* becomes known to the attackers. Fahad, the overlord security consultant, suggests that instead of regenerating the new keys completely from the scratch, only the new exponents *e', d'* need to be re-computed, leaving the modulus *n* unchanged (after all, indeed modulus computation requires more work).

Is this safe? If yes, explain why. If not, show how the pirates can compromise the new system (i.e. compute new *d'* from *e, d, n, e'*).

Attacker HKN intercepted some packets $c_1, c_2, \ldots, c_k$, encrypted using RSA with public exponent *e* and modulus $n=pq$ (i.e. HKN knows only ciphertexts, *n, e,* but not *p, q*). His spies also learnt that one of the plaintext packets m_i (for some *i*, s.t. $c_i = m_i^e \bmod n$) is divisible by *p*.
Can HKN decipher all intercepted packets now? How or why not?

We Will Make a Change

Nick and Nikhil use RSA cryptosystems with the same number n = pq, but different encryption exponents e_1 and e_2, which are relatively prime. Alex sends the message m to both Nick and Nikhil using their individual encryption keys. Show that Erik can recover the message m from the two ciphertexts.

RSA Solutions

Not safe: Fahad clearly has not taken this class.
We know that $e*d \bmod \varphi(n) = 1$.
Suppose, (using extended gcd) you found x such that
$x*e' \bmod (e*d-1) = 1$
Then $x*e' \bmod \varphi(n) = 1$, so $x=d'$.

Yes, HKN can now factor n as follows, and thus decipher not only the intercepted packets but also all the other messages encrypted with the same keys:
For each c_i compute $r_i = \gcd(c_i,n)$. If $1< r_i <n$, then $r_i = p$. This will be the case for i for which m_i is divisible by p. Indeed, if $m_i = p*k$, for some integer $k>0$, then $c_i=m_i^e \bmod n = p^e k^e \bmod (p*q)$.
Therefore, c_i is also divisible by p.
We do need to note two possible "failures" for HKN: if m_i which is divisible by p is either 0 or n. In either of those case HKN would learn nothing. But these two cases are "illegal" messages since neither message is in Z_n^*.

RSA Solutions

We Will Make a Change

Nick and Nikhil use RSA cryptosystems with the same number n = pq, but different encryption exponents e_B and e_C , which are relatively prime. Alex sends the message m to both Nick and Nikhil using their individual encryption keys. Show that Erik can recover the message m from the two ciphertexts.

$$1 = xe_B + ye_C$$

$$m = m^1 = m^{xe_B + ye_C} = (m^{e_B})^x \cdot (m^{e_C})^y = c_B^x \cdot c_C^y$$

Polynomials and ECC

1. We would like to send a message of length n over a channel. You know that at most k packets can be dropped, and of the packets not dropped, j might be corrupted. How many packets do we need to send in order to guarantee a correct decoding?
2. We have a polynomial of degree 2 that goes through the points (1,0), (2,3), and (4,0) modulo 7. What is P(3)?

Polynomials and ECC Solutions

1. N + k + 2j. The worst thing that can happen is k dropped (which is fixed by the + k) and then j get corrupted (but the other n + j non-corrupted packets will be enough to describe the overall polynomial). We get rid of both erasure errors as well as general errors.
2. Our polynomial is a(x-1)(x-4). We can solve for a by checking the value of the polynomial at x=2, we get a(1)(5) = 3 mod 7, and we get a is 2. So, we have 2(x-1)(x-4) Hence, P(3) is 2(2)(6) mod 7, which is 3.

Graphs

[True/False]: The hypercube graph always has an Eulerian tour

Let G be a non bipartite triangle-free simple graph with n vertices and minimum degree k. Let l be the minimum length of an odd cycle in G.

(a) Let C be a cycle of length l in G. Prove that every vertex not in $V(C)$ has at most two neighbors in $V(C)$.

(b) By counting the edges joining $V(C)$ and $V(G) - V(C)$ in two ways, prove that $n \geq kl/2$ (and thus $l \leq 2n/k$).

In a village there are three schools with n students in each of them. Every student from any of the schools is on speaking terms with at least $n + 1$ students from the other two schools. Show that we can find three students, no two from the same school, who are on speaking terms with each other

Let T1 and T2 be spanning trees of G with T1 != T2. Prove that there exists e that is in T1 and not T2 and f that is in T2 and not T1 so that both $T1 - e + f$ and $T2 - f + e$ are spanning trees.

Graphs Solutions

The hypercube graph always has an Eulerian tour False only true when dimension is even.

Let G be a nonbipartite triangle-free simple graph with n vertices and minimum degree k. Let l be the minimum length of an odd cycle in G.
(a)Let C be a cycle of length l in G. Prove that every vertex not in $V(C)$ has at most two neighbors in $V(C)$.
Suppose $x \in V(G) - V(C)$ has three neighbors in $V(C)$. Then those neighbors partition C into three pieces. Since l is odd, at least one of those parts has odd length. And since there are no triangles in G, that part has length at most $l - 4$. Then by walking from x to one of the endpoints of that part, along that part, and then back to x, you've built an odd cycle of length less than l, a contradiction.

(b) By counting the edges joining $V(C)$ and $V(G) - V(C)$ in two ways, prove that $n \geq kl/2$ (and thus $l \leq 2n/k$).
Since $\delta(G) = k$, the edges coming out of $V(C)$ is at least $kl - 2l$ (the sum of the degrees minus the edges contributing to C, which were all double-counted). On the other hand, the number of edges coming into $V(C)$ is at most $(|V(G)| - |V(C)|) * 2 = (n - l) * 2$. So $kl - 2l \leq (n - l) * 2 \implies kl \leq 2n \implies kl/2 \leq n$.

Graphs Solutions

In a village there are three schools with n students in each of them. Every student from any of the schools is on speaking terms with at least $n + 1$ students from the other two schools. Show that we can find three students, no two from the same school, who are on speaking terms with each other

Certainly, we are talking about a tripartite graph. Choose a vertex u adjacent with k vertices from another class (call it A) and at least $n + 1 - k$ vertices from the third class (call it B), **such that the k is the smallest possible**. Let v be a neighbor of u in A. Since v has at least k neighbors in B, it must have a common neighbor w with u. Then uvw is a triangle.

Let T1 and T2 be spanning trees of G with T1 != T2. Prove that there exists e that is in T1 and not T2 and f that is in T2 and not T1 so that both T1 −e+f and T2 −f +e are spanning trees.

Choose $e \in E(T1) \setminus E(T2)$ and let C be the fundamental cycle of e with respect to $T2$. Let u, v be the ends of e and observe that the subgraph $T1 - e$ has exactly two components, $H1$, $H2$ and we may assume (without loss) that $u \in V(H1)$ and $v \in V(H2)$. Now, $C - e$ is a path from u to v, so there must exist an edge $f \in E(C - e)$ so that f has one end in $H1$ and the other in $H2$. It now follows that both $T1 - e + f$ and $T2 + e - f$ are spanning trees, as desired.

Counting

1. There are 9 representatives in the UC Berkeley Computer Science Conference: 2 from 61A, 3 from 61B, and 4 from 70. During the opening ceremony, 3 of the representatives fall asleep. In how many ways can exactly two of the sleepers be from the same class?

2. Corrina is choosing fruit from a tray. She only wants 6 pieces of fruit and she can choose from grapes, strawberries, and raspberries. There are at least 6 pieces of each of these fruits on the tray in front of her. How many different assortments of 6 fruits can be selected?

3. There are 2000 white balls in a box, and unlimited supply of white, green and red balls, initially outside the box. in each turn, we may replace: 2 white with 1 green, 2 reds with green, 2 greens with white and red, a white and green with a red, or a green and red with a white. if we end up with just three balls, prove at least one ball is green.

Counting Solutions

There are 9 representatives in the UC Berkeley Computer Science Conference: 2 from 61A, 3 from 61B, and 4 from 70. During the opening ceremony, 3 of the representatives fall asleep. In how many ways can exactly two of the sleepers be from the same class?

61A: 1 * (9-2) = 7

61B: 3 * (9 - 3) = 18

70: 6 * (9 - 4) = 30

7 + 18 + 30 = 55

Corrina is choosing fruit from a tray. She only wants 6 pieces of fruit and she can choose from grapes, strawberries, and raspberries. There are at least 6 pieces of each of these fruits on the tray in front of her. How many different assortments of 6 fruits can be selected?

8 CHOOSE 2 or 7 CHOOSE 2 + 7

Counting Solutions

There are 2000 white balls in a box, and unlimited supply of white, green and red balls, initially outside the box. in each turn, we may replace: 2 white with 1 green, 2 reds with green, 2 greens with white and red, a white and green with a red, or a green and red with a white. if we end up with just three balls, prove at least one ball is green.

So we have 3 different cases
i)when we replace (2 white with 1 green or 2 reds with 1 green)
then $n \to (n-1)$ and $g \to (g+1)$
therefore $n-1 \equiv (g+1) \mod 2$

ii) When we replace 2 greens with white and red
$n \to n$ and $g \to (g+2)$
$n \equiv (g+2) \mod 2$
iii) When we replace (a white and green with a red, or a green and red with a white)
$n \to (n-1)$ and $g \to (g-1)$
then $(n-1) \equiv (g-1) \mod 2$

This means that if n is odd, g is also, so when $n = 3$ $g = 2k+1$ for $k \geq 0$ and at least one ball is green.

Halting Solutions

The function Neverloops(P) is 0 if program P does not halt on some input x, and 1 if P halts on every input x. Is there a program that computes Neverloops? Justify your answer

[True/False]: The problem of determining whether a program halts in time $2^{(n^2)}$ hours on an input of size n is undecidable.

Halting Solutions

The function Neverloops(P) is 0 if program P does not halt on some input x, and 1 if P halts on every input x. Is there a program that computes Neverloops? Justify your answer

Answer: No. Suppose that on the contrary, there is a program that computes Neverloops. Then we can use it to solve the halting problem through the following function:

```
def TestHalt(P, x):
  def P'(y):
    return P(x)
  if NeverLoops(P') == 1:
    return "halts"
  else:
    return "does not halt"
```

The function above returns "halt" if and only if program P on input x halts. This is the halting problem. So we see that we can easily solve the halting problem if we were able to define the program Neverloops. But we know that the halting problem is unsolvable. Hence, there is a contradiction. There is no program that computes Neverloops.

Halting Solutions

[True/**False**]: The problem of determining whether a program halts in time $2^{(n^2)}$ hours on an input of size n is undecidable.
Just run it

Probability

Recall that in a secret sharing scheme the secret $p(0) \bmod q$ can be reconstructed from the values of the polynomial $p(x)$ of degree d at any $d + 1$ points. However, the values of the polynomial $p(x)$ at any d points reveal absolutely no information about the secret $p(0)$. This condition can be formally stated using conditional probability as follows:

$$\Pr[p(0) = a \mid p(1), p(2), \ldots, p(d)] = 1/q \text{ for every } a \bmod q.$$

Now suppose Alice wishes to share a secret that consists of two numbers a and b, each mod q. She picks a random degree d polynomial $p(x) \bmod q$ such that $p(0) = a$ and $p(1) = b$. She distributes shares $p(2), \ldots, p(k)$ as with standard secret sharing (where $k \geq d + 2$), and claims that any $d + 1$ people can reconstruct the secret, but any d people have absolutely no information about the secret.

a. Formally state (using conditional probability) Alice's claim that the values $p(2), \ldots, p(d + 1)$ reveal absolutely no information about the secret a, b.

b. Is Alice's claim correct? If so prove it, if not give a precise reason why not.

Probability Solutions

a)

$$\Pr[p(0) = a, p(1) = b \mid p(2), p(3), \ldots, p(d+1)] = \frac{1}{q^2}.$$

Common Mistakes:

- Set the above equation to $\frac{1}{q}$. Note that there are q possible values for a and b independently, for a total of q^2 values for the secret.
- Wrote separate equations for $\Pr[p(0) = a]$ and $\Pr[p(1) = b]$, setting them both to $\frac{1}{q}$. While this is a correct equation, it doesn't claim anything about the independence of a and b, which is required for all possible values of the secret to be equally likely. (This statement is in fact true.)

Answer: No, because

b)

$$\Pr[p(0) = a, p(1) = b \mid p(2), p(3), \ldots, p(d+1)] = \frac{1}{q}$$

if the $d+2$ points $p(0) = a, p(1) = b, p(2), \ldots, p(d+1)$ lie on a degree d polynomial, and equal to 0 otherwise. This is because $d+2$ points do not necessarily lie on a degree d polynomial, so some (a, b) pairs are invalid. Alternatively, you could say that knowing one of the values a, b uniquely determines the other, or that there are only q possible polynomials with the points given, but there should be q^2 possible values for (a, b).

Countability

1. Give a bijection from the real number interval (1,oo) to the real number interval (0,1). What does it say about both of these sets?
2. Prove that the set of all programs if countably infinite.

Countability Solutions

1. 1/x
2. All programs can be represented as a finite bitstring over a language, therefore, as the set of all finite bit strings is countable, so is the set of all programs.

Variance/Expectation/Distributions

Anwar, Nikhil

Properties of Expectation and Variance

$E(X + Y) = E(X) + E(Y)$

$E(cX) = cE(X)$

$E(XY) = E(X)E(Y)$ if X and Y independent

$Var(X + Y) = Var(x) + Var(Y)$ if X and Y independent

$Var(cX) = c^2Var(X)$

Conditional Expectation

If a coin is biased, it will show up heads 80% of the time. There is a 20% chance that it is biased. What is the expected number of heads we should get when we flip this same coin 100 times.

Answer

X = number of heads

$E[X|\text{biased}] = 80$
$E[X|\text{fair}] = 50$

$Pr[\text{biased}] = 0.2$
$Pr[\text{fair}] = 0.8$

$$E[X] = E[X|\text{biased}] * Pr[\text{biased}] + E[X|\text{fair}] * Pr[\text{fair}]$$

$$\begin{aligned} E[X] &= 80*0.2 + 50*0.8 \\ &= 16 + 40 \\ &= 56 \end{aligned}$$

Expectation & Variance Practice

You roll a die twice. What is the expectation of the minimum value between the two faces? The variance?

Answer:

1: 11/36 (if either is 1: 6 + 5, account for overlap)
2: 9/36
3: 7/36
4: 5/36
5: 3/36
6: 1/36 (only if both are 1)

E[x] = 1*(11/36) + 2*(9/36) + 3*(7/36)
+ 4*(5/36) + 5*(3/36) + 6*(1/36)

E[x] = (11+18+21+20+15+6)/36
= 91/36 = 2.53

E[x^2] = 1*(11/36) + 4*(9/36) + 9*(7/36) + 16*(5/36) + 25*(3/36) + 36*(1/36)
= (11+36+63+80+75+36)/36
= 301/36 = 8.361

Var[x] = E[x^2] - (E[x])^2
= (301/36) - (91/36)^2
= 1.971

Distributions to know

Name	Pr[X = k]	E(X)	Var(X)
Bernouli(p)	1: p, 0: 1-p	p	p(1-p)
Uniform(n)	1/n	(n+1)/2	$(n^2 - 1)/12$
Binomial(n, p)	(n choose k)*$p^k(1-p)^{n-k}$	np	np(1-p)
Geometric(p)	$(1-p)^{k-1}p$	1/p	$(1-p)/p^2$
Poisson(**λ**)	$(\lambda^k e^{-\lambda})/k!$	λ	λ

Which Distribution??

Number of days until you begin studying for CS70 (you start studying with chance p)

Number of days you'll waste procrastinating before your final (you procrastinate with chance p)

Number of times that the teacher will give everyone a clarification during the exam (at any moment, can happen with chance p)

Which room you take your final in, chance that you are taking your final in 1 Pimentel (out of 2 possible rooms) is p

Answer:

Number of days until you begin studying for CS70 Geometric

Number of days you'll waste procrastinating before your final Binomial

Number of times that the teacher will give everyone a clarification during the exam Poisson

Chance that you are taking your final in [insert room here] Bernouli

Markov and Chebyshev

Albert

Markov's Inequality

$$Pr[X \geq \alpha] \leq \frac{E[X]}{\alpha}^{*}$$

*α ≥ 0,
∀α

Mark off Madoff

It's 2008 and Bernie Madoff is selling pipe dreams. You're smart some days and impulsive on others. 20% of the time, you're smart and don't invest that day. The rest of the time, you let temptation get the better of you and the amount you invest (lose) is distributed over a Poisson with variance $2 million.

After nearly a month (25 days), Madoff is arrested.

Bound the probability of losing more than a $100 million.

Mark off Madoff Solution

Let X be a random variable that denotes how much you lose. Let X_i be how much you lose on the i^{th} day.

$$X_i \sim \begin{cases} Pois(\lambda), Pr[X_i \sim Pois(\lambda)] = 0.8 \\ 0, Pr[X_i \sim 0] = 0.2 \end{cases}$$

$$\lambda = E[X_i | X_i \sim Pois(\lambda)] = Var[X_i | X_i \sim Pois(\lambda)] = 2$$

$$E[X] = \sum_{i=1}^{25} E[X_i]$$

$$= 25(0.2E[X_i | X_i \sim 0] + 0.8E[X_i | X_i \sim Pois(2)])$$

$$= 25(0.8 * 2) = 40$$

$$Pr[X \geq 100] \leq \frac{40}{100} = 40\%$$

Chebyshev's Inequality

$$Pr[|X - \mu| \geq \alpha] \leq \frac{Var[X]}{\alpha^2}$$

What do you call a fat mathematician?

Chubbyshev.

I’m not sorry.

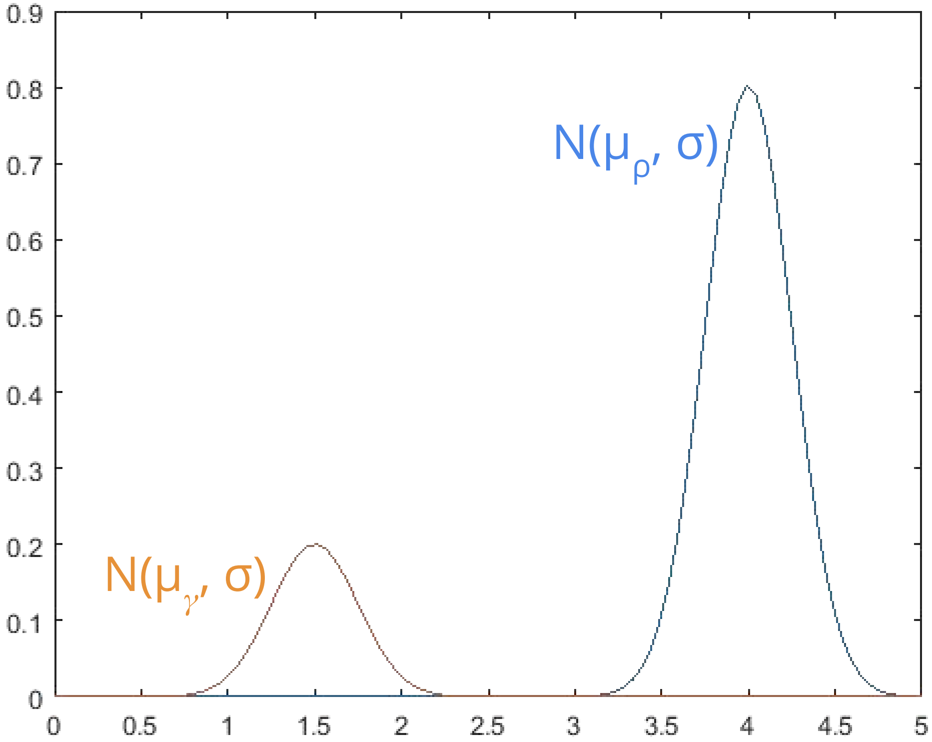
Fake it 'til you make it

You're pitching your latest app (Tinder for playdates for dogs) to a local VC.

As any CEO knows, the amount that the VC will invest depends on your level of confidence when pitching, which depends on how prepared you are. Say you prepare a week ahead with probability 80% and the night before the rest of the time, resulting in high and low confidence, respectively.

After surveying the data, you learn that high-confidence pitches average $4 million while low-confidence pitches net $1.5 million on average, both with standard deviation of $0.25 million, distributed normally.

Use Chebyshev's to find a 98% confidence interval (2 sig figs).



$\mu_\gamma = 1.5$, $\mu_\rho = 4$, $\sigma = .25$

$\Pr[X \in (\text{lower}, \text{upper})] = 0.98$

Fake it 'til you make it Solution

Let α be the area under the N(μ_γ, σ); let β be the area under the N(μ_ρ, σ). Then α+β is the total area under the distribution. We seek an interval that contains 0.98α+0.98β.

We find the left bound of a 96% CI for N(μ_γ, σ) and the right bound of a 96% CI for N(μ_ρ, σ)

$$Pr[|X - \mu_\rho| > \alpha] = Pr[|X - \mu_\gamma| > \alpha] \leq \frac{0.25^2}{\alpha^2}$$

$$\frac{0.25^2}{\alpha^2} = 0.04$$

$$\alpha = 1.25$$

$$CI_\gamma = [\mu_\gamma - 1.25, \mu_\gamma + 1.25] = [1.5 - 1.25, 1.5 + 1.25] = [0.25, 2.75]$$

$$CI_\rho = [\mu_\rho - 1.25, \mu_\rho + 1.25] = [4 - 1.25, 4 + 1.25] = [3.75, 5.25]$$

$$CI_{\gamma+\rho} = [0.25, 5.25]$$

Bonus! What's the mean of the distribution? $\mu_{\gamma+\rho} = 0.8\mu_\gamma + 0.2\mu_\rho$

Covariance

Erik

Covariance

Cov(X, Y), the covariance of the random variables X and Y, is a measure of how much X and Y change together.

$$\begin{aligned} \mathrm{Cov}(X, Y) &:= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[X]E[Y]. \end{aligned}$$

With a positive covariance, when X or Y increases, we expect the other to also increase. With a negative covariance, when one increases, we expect the other to decrease.

To calculate, we need to find E(X), E(Y), and E(XY)

Properties of Covariance

From lecture:

1. $var[X] = cov(X,X)$
2. X,Y independent $\Rightarrow cov(X,Y) = 0$
3. $cov(a+X,b+Y) = cov(X,Y)$
4. $cov(aX + bY, cU + dV)$

   $= ac*Cov(X,U) + ad*Cov(X,V) + bc*Cov(Y,U) + bd*Cov(Y,V)$

Other Properties of Covariance

If time, prove:

1. $Cov(X, c) = 0$ for X: any random variable, c: constant
2. $(Cov(X, Y) = 0) \nRightarrow$ X, Y independent
3. $Cov(\sum_i(a_i X_i) , \sum_j(b_j Y_j)) = \sum_i \sum_j(a_i b_j * Cov(X_i, Y_j))$ for a_i, b_j constants, X_i, Y_j random variables, $i \in (1, n)$, $j \in (1, m)$, n, m: natural numbers

Calculating Covariance

1. You flip coin c_1 twice and it lands heads with probability $p_1 = .5$. Every time it lands heads, you flip another coin c_2 that lands heads with probability $p_2 = .5$. What is the covariance between the number of times c_1 lands heads and the number of times c_2 lands heads?
2. X, Y, Z are i.i.d. N(0,1)
   $A = X + Y + Z$
   $B = fX + gY + hZ$ for some constants f, g, h
   Find Cov(A, B).

   i.i.d. : independent and identically distributed
   N(a, b): normal distribution with mean a, variance b
3. Continue 2: $C = sA + tB$, find Cov(A, C)

Calculating Covariance Answers (1)

1. X:= number of times c_1 lands heads, Y:= number of times c_2 lands heads
   $E(X) = 2*p_1 = 1$, $E(Y) = 2*p_1p_2 = ½$
   $E(XY) = \sum_{a \in A} a*Pr(XY = a)$, $A=\{0, 1, 2, 4\}$
   $0*Pr(XY=0) = 0$
   $Pr(XY = 1) = Pr(X=1, Y=1) = Pr(X=1)*Pr(Y=1|X=1) = 2*p_1*(1-p_1)*p_2 = ¼$
   $Pr(XY = 2) = Pr(X=2, Y=1) = Pr(X=2)*Pr(Y=1|X=2) = p_1^2*2*p_2*(1-p_2) = ⅛$
   $Pr(XY = 4) = Pr(X=2, Y=2) = Pr(X=2)*Pr(Y=2|X=2) = p_1^2*p_2^2 = 1/16$
   $E(XY) = 0 + 1*(¼) + 2*(⅛) + 4*(1/16) = ¾$
   $Cov(X,Y) = E(XY)-E(X)E(Y) = ¾ - (1*½) = ¼$

Calculating Covariance Answers (2)

2. X, Y, Z i.i.d. N(0,1), A = X + Y + Z, B = fX +gY + hZ

Could try: Cov(A, B) = E(AB) - E(A)E(B) *hard to find E(AB), so instead

We use the third fact from Other Properties of Covariance slide:

$$\mathrm{Cov}\left(\sum_i (a_i X_i) , \sum_j (b_j Y_j) \right) = \sum_i \sum_j \left(a_i b_j * \mathrm{Cov}(X_i, Y_j) \right)$$

Cov(A, B) = Cov(X+Y+Z, fX+gY+hZ)

= f*Cov(X, X) + g*Cov(X, Y) + h*Cov(X,Z) + f*Cov(Y, X) + g*Cov(Y,Y) + h*Cov(Y, Z) + f*Cov(Z, X) + g*Cov(Z,Y) + h*Cov(Z,Z)

(X, Y, Z independent, so)

= f*Cov(X, X) + g*Cov(Y, Y) + h*Cov(Z,Z) = f*Var(X) + g*Var(Y) + h*Var(Z)

= f+g+h

Calculating Covariance Answers (3)

3. C = sA + tB, find Cov(A, C)

Cov(A, C) = Cov(A, sA + tB)

use the third fact from Other Properties of Covariance slide:

$Cov(\sum_i(a_iX_i) , \sum_j(b_jY_j)) = \sum_i\sum_j(a_ib_j*Cov(X_i,Y_j))$

= s*Cov(A, A) + t*Cov(A, B)

= s*Var(A) + t*(f+g+h)

Var(A) = Var(X+Y+Z) *X,Y,Z independent so* = Var(X) + Var(Y) + Var(Z) = 3

= s*3 + t*(f+g+h)

LLSE
(Linear Least Square Estimate)

Nick, Corrina

LLSE: Motivation

- Common statistical problem: given a variable Y that depends on another variable X, how to predict Y given value of X?
- For simplicity of computation, want to estimate Y with a *linear* function of X
- Say we have two random variables **X** and **Y**, and we know their joint probability distribution ($\Pr[X = x, Y = y]$ for all values x, y)
- Want the *best* estimate of Y as a linear function of X

LLSE: Definition

$$L(Y|X) := E(Y) + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}(X - E(X))$$

L(Y|X) is a random variable whose value depends on that of **X**

LLSE: Definition

$$L(Y|X) := \boxed{E(Y)} + \frac{\boxed{\mathrm{Cov}(X,Y)}}{\boxed{\mathrm{Var}(X)}}(X - \boxed{E(X)})$$

constants

note that **X** is the only random variable in the definition of **L(Y|X)**

LLSE: Analysis

Theorem: Least Squares Property

L(Y|X) is the linear function of **X** with the least squared difference from **Y**

$$E\left((Y - L(Y|X))^2\right) \le E\left((Y - aX - b)^2\right)$$

for any function **aX+b**

credit, and for more details: Sinho's awesome notes, bit.ly/sinho_notes

LLSE: Analysis

Theorem: Least Squares Property

L(Y|X) is the linear function of **X** with the least squared difference from **Y**

$$\underbrace{E\left((Y - L(Y|X))^2\right)}_{\text{least-squares difference for LLSE}} \leq \underbrace{E\left((Y - aX - b)^2\right)}_{\text{least-squares difference for } \mathbf{aX+b}}$$

for any function **aX+b**

this is why we care about the LLSE!

credit, and for more details: Sinho's awesome notes, bit.ly/sinho_notes

LLSE: Simple Practice Problems

0. **X**, **Y** independent

1. **Y** = 2**X** w.p. ½, 1 w.p. ½; **var(X)** = 8

*What is **L(Y|X)**?*

2. **X** = die roll

Y = payout: 3 * that die roll + 2 * a coin toss

*What is **L(Y|X)**?*

LLSE: Simple Practice Problems

0. **X**, **Y** independent

1. **Y** = 2**X** w.p. ½, 1 w.p. ½; **var(X)** = 8

   *What is **L(Y|X)**?*

2. **X** = die roll

   Y = payout: 3 * that die roll + 2 * a coin toss

   *What is **L(Y|X)**?*

LLSE: Simple Practice Problems *solutions*

0. **cov(X, Y)** = 0, so **L(Y|X)** = E[Y]

LLSE: Simple Practice Problems *solutions*

1. $E[Y] = \frac{1}{2} E[2X] + \frac{1}{2} E[1] = E[X] + \frac{1}{2}$

   $E[XY] = \frac{1}{2} E[2X^2] + \frac{1}{2} E[1X] = E[X^2] + \frac{1}{2} E[X]$

   $\mathrm{cov}(X, Y) = E[XY] - E[X]\, E[Y]$

   $= E[X^2] + \frac{1}{2} E[X] - E[X]\, (E[X] + \frac{1}{2})$

   $= E[X^2] + \frac{1}{2} E[X] - E[X]^2 - \frac{1}{2} E[X]$

   $= E[X^2] - E[X]^2$

   $= \mathrm{var}(X) = 8$

LLSE: Simple Practice Problems *solutions*

1, continued.

$$\mathbf{L(Y|X)} = \mathbf{E[Y]} + \frac{8}{8}(\mathbf{X} - \mathbf{E[X]})$$

$$= \mathbf{E[Y]} + \mathbf{X} - \mathbf{E[X]}$$

$$= (\mathbf{E[X]} + \tfrac{1}{2}) + \mathbf{X} - \mathbf{E[X]}$$

$$= \mathbf{X} + \tfrac{1}{2}$$

LLSE: Simple Practice Problems *solutions*

2. **X** = die roll

   Y = payout: 3 * that die roll + 2 * a coin toss

   *What is **L(Y|X)**?*

LLSE: Simple Practice Problems *solutions*

2. **X** = die roll

Y = payout: 3 * that die roll + 2 * a coin toss

*What is **L(Y|X)**?*

*Let **C** be the result of the coin flip*
***E[C]** = ½*

Y = 3 **X** + 2 **C**

LLSE: Simple Practice Problems *solutions*

$E[Y] = E[3X + 2C]$
$= 3\,E[X] + 2\,E[C]$
$= 3\,(7/2) + 2\,(1/2)$
$= 23/2$

2 (cont).

$$\begin{aligned}
\mathrm{cov}(X, Y) &= E[XY] - E[X]\,E[Y] \\
&= E[3X^2 + 2CX] - E[X]\,E[Y] \\
&= 3\,E[X^2] + 2\,E[C]\,E[X] \quad - \quad (7/2)\,(23/2) \\
&= 3\,(1/6)\,(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) + 2\,(\tfrac{1}{2})\,E[X] \quad - \quad (161/4) \\
&= (1/2)\,(1 + 4 + 9 + 16 + 25 + 36) + (7/2) - (161/4) \\
&= (91/2) + (7/2) - (161/4) \\
&= 35/4
\end{aligned}$$

*(because **C** and **X** are independent)*

LLSE: Simple Practice Problems *solutions*

2 (cont). $$L(Y|X) = E[Y] + \frac{\text{cov}(X, Y)}{\text{var}(X)} (X - E[X])$$

variance of a die roll is (35 / 12)

$$= (23 / 2) + \frac{(35 / 4)}{(35 / 12)} (\mathbf{X} - (7 / 2))$$

$$= \boxed{(23 / 2) + 3 (\mathbf{X} - (7 / 2))}$$

LLSE Practice Problem

- Y describes the temperature and X describes the probability of rain
- X and Y are inversely proportional
- Example: Given a colder temperature, our prediction of rain should increase

$$L[X|Y] = a + bY = E(X) + \frac{\mathrm{cov}(X,Y)}{\mathrm{var}(Y)}(Y - E(Y))$$

LLSE Practice Problem

- Cov(X, Y) will be negative! A higher temperature has lower chance of rain
- The slope changes our prediction of X when given Y's information

$$L[X|Y] = a + bY = E(X) + \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)}(Y - E(Y))$$

LLSE Practice Problem

Jean is practicing his prediction skills. The percentage of clouds Y, is a uniform r.v between 0 and 1. The temperature Z, is a uniform r.v between 0 and 1. The probability of rain X, is given by the average of Y and Z. Unfortunately Jean only has access to a machine that measures the value Y+2Z.

$$L[X|Y+2Z] = E[X] + \frac{Cov(X, Y+2Z)}{Var(Y+2Z)}(Y + 2Z - E[Y+2Z])$$

LLSE Practice Problem

Define the random variables

$$Y = Uniform(0,1)$$

$$Z = Uniform(0,1)$$

$$X = \frac{Y+Z}{2}$$

$$L[X|Y+2Z] = E[X] + \frac{Cov(X, Y+2Z)}{Var(Y+2Z)}(Y+2Z-E[Y+2Z])$$

LLSE Practice Problem

Calculate each part of the expression. Start with the easiest ones.

$$E[X] = \frac{E[Y] + E[Z]}{2} = \frac{0.5 + 0.5}{2} = 0.5$$

$$Var(Y + 2Z) = Var(Y) + Var(2Z) = \frac{1}{12} + 4\frac{1}{12} = \frac{5}{12}$$

$$E[Y + 2Z] = 1.5$$

LLSE Practice Problem

Compute the covariance term.

$$Cov(X, Y+2Z) = Cov(\frac{Y+Z}{2}, Y+2Z)$$

Simplify covariance into smaller problems

$$Cov(\frac{Y}{2}, Y) + Cov(\frac{Y}{2}, 2Z) + Cov(\frac{Z}{2}, Y) + Cov(\frac{Z}{2}, 2Z)$$

LLSE Practice Problem

Compute the covariance term

$$Cov(X, Y+2Z) = Cov(\frac{Y+Z}{2}, Y+2Z)$$

$$Cov(\frac{Y}{2}, Y) + Cov(\frac{Y}{2}, 2Z) + Cov(\frac{Z}{2}, Y) + Cov(\frac{Z}{2}, 2Z)$$

$$\frac{1}{2}Cov(Y, Y) + 0 + 0 + Cov(Z, Z)$$

LLSE Practice Problem

Covariance between two independent r.v. is 0. Thus we can set Cov(Y,Z) to 0 since they're independent.

$$Cov(\frac{Y}{2}, Y) + Cov(\frac{Y}{2}, 2Z) + Cov(\frac{Z}{2}, Y) + Cov(\frac{Z}{2}, 2Z)$$

$$\frac{1}{2}Cov(Y, Y) + 0 + 0 + Cov(Z, Z)$$

LLSE Practice Problem

Use properties of covariance to solve everything in terms of variance

$$\frac{Var(Y)}{2} = \frac{1}{24}$$

$$Var(Z) = \frac{1}{12}$$

$$Cov(X, Y + 2Z) = \frac{1}{24} + \frac{1}{12} = \frac{1}{8}$$

LLSE Practice Problem

Solve the slope of the relationship by dividing the variance of the Y + 2Z

$$\frac{Cov(\frac{Y+Z}{2}, Y+2Z)}{Var(Y+2Z)} = \frac{\frac{1}{8}}{\frac{5}{12}} = \frac{3}{10}$$

LLSE Practice Problem

Plug all the expressions and solve for (a, b)

$$L[X|Y+2Z] = E[X] + \frac{Cov(X, Y+2Z)}{Var(Y+2Z)}(Y+2Z-E[Y+2Z])$$

$$L[X|Y+2Z] = \frac{1}{2} + \frac{3}{10}(Y+2Z-1.5) = 0.05 + \frac{3(Y+2Z)}{10}$$

$$(a, b) = (\frac{3}{10}, 0.05)$$

Continuous Distributions

Alex T/Fahad

Continuous vs. Discrete

	Continuous	Discrete
Defining a distribution	Density function f(x)	Dictionary of assignments {x: p1, y: p2,...}
Expectation	Weighted integral	Weighted sum
Variance	$E[X^2] - E[X]^2$	$E[X^2] - E[X]^2$

Distributions to know

<u>Discrete</u>	<u>Continuous</u>
Uniform - Equal probability of each element - Question: How can we get a subset of these elements to show up?	**Uniform** - Equal PDF value within some single finite interval - Question: How can we get an interval of these values to show up?
No analog* * There exists a discrete normal distribution; applications of it are identical to the continuous normal distribution, but it is often used as an approximation	**Normal/Gaussian** - Infinite support (region of non-zero PDF) - Symmetric about the mean - Random samples from any population - Question: What proportion of my data is this far away from the mean?

Distributions to know

Discrete	Continuous
Binomial - Discrete trials (n) - Independent and identical probability of success in each trial (p) - Question: how many successes in n trials? **Poisson*** - Infinite, continuous trials ($n \to \infty$) - Very small probability of success ($p \to 0$) - Defined by average number of successes in an interval of continuous trials ($np = \lambda$) - Question: how many successes in some interval of continuous trials? * Poisson is considered a *discrete* distribution because the question is asking about something discrete, despite the "trials" being continuous	

Distributions to know

Discrete	Continuous
Geometric - Discrete trials (n) - Independent and identical probability of success in each trial (p) - Question: how many trials until the first success?	**Exponential** - Infinite, continuous trials ($n \rightarrow \infty$) - Very small probability of success ($p \rightarrow 0$) - Defined by average number of successes in an interval of continuous trials ($np = \lambda$) - Question: how long to wait until the first success?

Poisson Distribution

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

$P(X = k) = e^{-\lambda}\lambda^k/k!$

$E[X] = \lambda$

$\text{Var}[X] = \lambda$

e.g. inst.eecs.berkeley.edu fails twice per month, on average. Assuming failures are independent, what is the probability that the server will fail at least once in these last 2 weeks, when the 61B project is due?

Exponential Distribution

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

$f(x) = e^{-\lambda x}\lambda$ if $x >= 0$; 0 if $x < 0$;

$E[X] = 1/\lambda$

$Var[X] = 1/\lambda^2$

e.g. inst.eecs.berkeley.edu fails twice per month, on average. Assuming failures are independent, what is the probability that once school starts, we go an entire month until it fails, when the first 61B project is released?

Gaussian Distribution

$X \sim \text{Normal}(\mu, \sigma)$

$f(x) = (1/\sigma\sqrt{(2\pi)})\ e^{-(x-\mu)^{\wedge}2/(2\sigma)^{\wedge}2}$

$E[X] = \mu$

$Var[X] = \sigma^2$

e.g. In the overall human population, the average IQ is 100 with a standard deviation of 15. Katya's IQ is 55. What is the probability that a randomly selected individual is at least as smart as Katya?

Which distribution and which value?

Due to human impact on climate and habitats, fewer and fewer Roan antelope can be seen in South Africa’s Kruger National Park. If there are an average of 2 sightings of Roan antelope per day, what is the probability that there are 7 sightings in a week?

Which distribution and which value?

Due to human impact on climate and habitats, fewer and fewer Roan antelope can be seen in South Africa's Kruger National Park. If there are an average of 2 sightings of Roan antelope per day, what is the probability that there are 7 sightings in a week?

Poisson.
Since our $\lambda = 2$ per day, we take our query to be in terms of days.
Seeing 7 antelope per week is 1 per day (on average).
$P(X = 1) = \lambda e^{-\lambda}/1! = 2e^{-2} \approx 0.27$

Which distribution and which value?

The average score on a recent exam was 70%. Only 10 students scored above 90%. How many students are expected to have scored below 50%?

Which distribution and which value?

The average score on a recent exam was 70%. Only 10 students scored above 90%. How many students are expected to have scored below 50%?

Gaussian.
Ideally, exam scores are distributed normally. Gaussian distributions are symmetric about the mean (70% in this case), so the number of students who score 20% above the mean is expected to be the same as the number of students who score 20% below the mean.

Which distribution and which value?

On average, there are 6 independent attempts to pull a fire alarm during a single 61A exam that takes 3 hours. Assuming that the first attempt will always succeed, how far into the exam do we expect to go, until the fire alarm rings?

Which distribution and which value?

On average, there are 6 independent attempts to pull a fire alarm during a single 61A exam that takes 3 hours. Assuming that the first attempt will always succeed, how far into the exam do we expect to go, until the fire alarm rings?

Exponential.
Here, λ = 2 attempts per hour.
The question we're asking is time until the first attempt.
If $X \sim \text{Exponential}(\lambda)$, then $E[X] = 1/\lambda = ½$ hour.

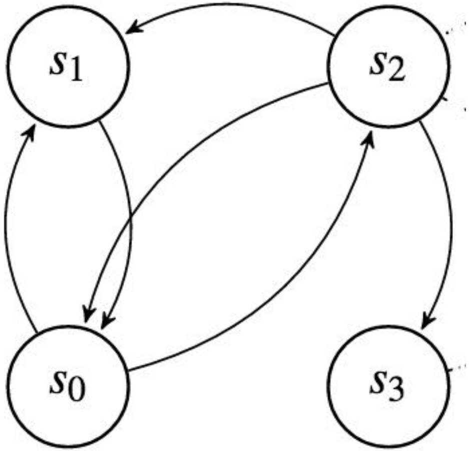
Markov Chains

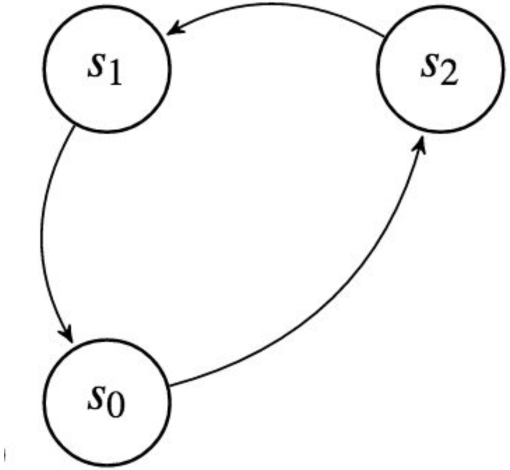
Michael

Some Definitions

Markov Chains consists of states and transition probabilities.

A Markov Chain is **irreducible** if you go from one state to any other state (including itself). Otherwise, it is **reducible**.





Periodicity

The period of any state in the graph is the same.

Let's define a set R(i) for every state.

$$\forall n \in N, P(X_n = i | X_0 = i) > 0 \Leftrightarrow n \in R(i).$$

This set contains n if we can take a sequence of n arrows from state i to state i.

If GCD(R(i)) = 1, a Markov Chain is **aperiodic**. Otherwise, it is **periodic**.

Why do these definitions matter?

An irreducible Markov Chain has an unique invariant distribution π ($\pi P = \pi$) .

The long term fraction of time spent in each state approaches π.

If a Markov Chain is also aperiodic, the distribution π_n converges to π as n grows large.

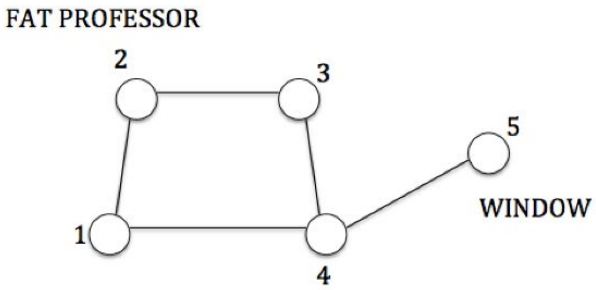


Figure 2: A fly wanders randomly on a graph.

(a) Suppose that the fly wanders as follows: if it is at node i at time n, then it chooses one of its neighbors j of i uniformly at random, and then wanders to node j at time $n+1$. For times $n = 0, 1, 2, \ldots$, let X_n be the fly's position at time n. Argue that $\{X_n, n \geq 0\}$ is a Markov chain, and find the invariant distribution.

(a) Suppose that the fly wanders as follows: if it is at node i at time n, then it chooses one of its neighbors j of i uniformly at random, and then wanders to node j at time $n+1$. For times $n = 0, 1, 2, \ldots$, let X_n be the fly's position at time n. Argue that $\{X_n, n \geq 0\}$ is a Markov chain, and find the invariant distribution.

Solution: Given the position of the fly at time n, the distribution of the position of the fly at time $n+1$ is conditional independent of the previous positions of the fly before n. Therefore, $\{X_n, n \geq 0\}$ is a Markov chain. We can get the probability transition matrix

$$P = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

According to $\pi P = \pi$, we get the invariant distribution $\pi = [0.2 \;\; 0.2 \;\; 0.2 \;\; 0.3 \;\; 0.1]$.

(b) Now for the process in part (a), suppose that the (not-to-be-named) professor sits at node 2 reading a heavy book. The professor is very fat, so he/she doesn't move at all, but will drop the book on the fly if it reaches node 2 (killing it instantly). On the other hand, node 5 is a window that lets the fly escape. What is the probability that the fly escapes through the window supposing that it starts at node 1?

(b) Now for the process in part (a), suppose that the (not-to-be-named) professor sits at node 2 reading a heavy book. The professor is very fat, so he/she doesn't move at all, but will drop the book on the fly if it reaches node 2 (killing it instantly). On the other hand, node 5 is a window that lets the fly escape. What is the probability that the fly escapes through the window supposing that it starts at node 1?

Solution: Let p be the probability that the fly escapes through the window supposing that it starts at node 1. According to symmetry, starting from node 3, the probability that the fly escapes through the window is also p. Let q be the probability that the fly escapes through the window supposing that it starts at node 4. We have

$$p = \frac{1}{2}(0+q)$$
$$q = \frac{1}{3}(1+p+p)$$

Then we get $p = \frac{1}{4}$.

(c) Now suppose that the fly wanders as follows: when it is at node i at time n, it chooses uniformly from all neighbors of node i except for the one that it just came from. For times $n = 0, 1, 2, \ldots$, let Y_n be the fly's position at time n. Is this new process $\{Y_n, n \geq 0\}$ a Markov chain? If it is, write down the probability transition matrix; if not, explain why it does not satisfy the definition of Markov chains.

(c) Now suppose that the fly wanders as follows: when it is at node i at time n, it chooses uniformly from all neighbors of node i except for the one that it just came from. For times $n = 0, 1, 2, \ldots$, let Y_n be the fly's position at time n. Is this new process $\{Y_n, n \geq 0\}$ a Markov chain? If it is, write down the probability transition matrix; if not, explain why it does not satisfy the definition of Markov chains.

Solution: No, it is not a Markov chain. According to the definition of the process

$$P(Y_{n+1} = 1 | Y_n = 4, Y_{n-1} = 1) = 0,$$

while

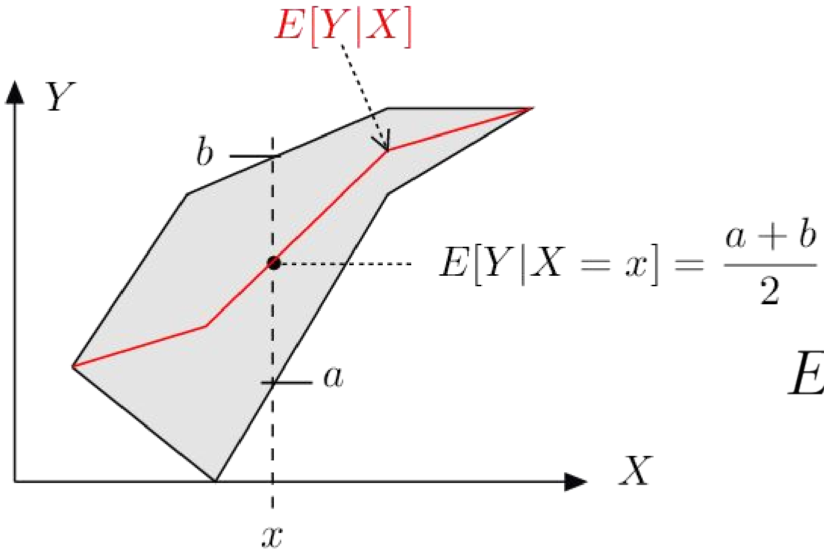
$$P(Y_{n+1} = 1 | Y_n = 4, Y_{n-1} = 3) = 0.5.$$

Therefore, given Y_n, Y_{n+1} and Y_{n-1} are not conditional independent. Then the process $\{Y_n, n \geq 0\}$ is not a Markov chain.

Conditional Expectation

Albert

Conditional expectation



$$E[Y|X=x] = \sum_y y Pr[Y=y|X=x]$$

$$E[Y|X=x] = g(y) : \Omega \to \mathbb{R}$$

$$E[Y|X] = \sum_x E[Y|X=x] Pr[X=x]$$

$$E[Y|X] = f(x) : \Omega \to g(y)$$

Condition your expectations

You're hoping to get a high score on the CS70 final, but not surprisingly, the number of questions you get wrong depends on how much you study. Say you study in 1 hour sessions, where the probability of the number of sessions, S, you study decreases linearly such that $\Pr[S=s] = (2/576)(24-s)$.

The number of questions Q you get wrong is distributed over Binom(100, p), where $p=\frac{1}{2}\Pr[S=s]$.

Find $E[Q|S=12]$ and $E[Q|S]$.

Condition your expectations solution

$$E[Q|S=12] = E[Q|Q \sim Binom(100, \frac{Pr[S=12]}{2})]$$
$$= 100 * 0.02 = 2$$

$$Pr[S=12] = \frac{2(12)}{576} \approx 0.04$$

$$E[Q|S] = E[Q|Q \sim Binom(100, \frac{Pr[S=s]}{2})]$$
$$= 100 * \frac{4(24-s)}{576} = \frac{25(24-s)}{36}$$