Midterm 1

Note 0: Review of Sets and Notations

cardinality: size of the set

cartesian product: set of all pairs with corresponding element.

Note 1: Logic

proposition: either true or false

DeMorgan's Law: $\neg(P \land Q) \equiv \neg P \lor \neg Q$

$\neg(P \lor Q) \equiv \neg P \land \neg Q$

Test Logic

- $P \to q \equiv \neg P \lor q$

Well Ordering Principle

- $(\forall n \in \mathbb{N}, \neg P(n+1) \Rightarrow \neg P(n)) \Rightarrow (\neg P(0) \lor \forall n \in \mathbb{N}, P(n))$
- $(\neg P(0) \land \forall n \in \mathbb{N}, (P(n) \Rightarrow P(n+1))) \equiv \forall n \in \mathbb{N}, \neg P(n)$

Distribution of Negation:

- $\neg(\forall x) P(x) \equiv (\exists x) \neg P(x)$
- $\neg \exists x P(x) \equiv \forall x \neg P(x)$
- $\neg(\forall x \exists y P(x,y)) \equiv \exists x \forall y \neg P(x,y)$
- $\forall x \exists y [P(x) \lor Q(y)] \equiv [\forall x P(x)] \lor [\exists y Q(y)]$

P and q are independent so can move quantifiers inside.

Note 3. Induction

1. Base case
2. Induction Hypothesis
3. Induction Step

$a | b \Rightarrow b = qa$

Note 4: The Stable Marriage Problem

SMAlgorithm

1) Every man proposes to the most preferred woman on his list who has not yet rejected him
2) Each woman collects all the proposals and puts on a string the man she likes best.
3) Each rejected man crosses off the woman that rejected him from his list.

- Repeat until each woman has a man on a string*

Lemma 1: The SMA always halts

Rogue Couple: man and woman that prefer each other over their current couple.

stable pairing: if no rogue couple

Lemma 2: (Improvement Lemma): if man M proposes to woman W on the k^{th} day, then on every subsequent day, W has on a string she likes at least as much as M.

Well Ordering Principle: If $S \subseteq \mathbb{N}$ and $S \neq \emptyset$, then S has a smallest element.

Lemma 3: The SMA always ends w/ a pairing that is stable

Theorem 4.2: SMA is male optimal,

Theorem 4.3: if male optimal $\Rightarrow$ female pessimal.

- if man prefers pairing S, then woman prefers T.

Note 5: Graph Theory

path: walk with no repeat vertex

walk: sequence of edges

tour: walk with no repeat edges

cycle: walk with no repeat vertex and $v_0 = v_f$

Euler's Tour: tour visiting every edge

Euler's theorem: undirected graph has Euler Tour iff G is even degree & connected

Hamiltonian Tour: tour visiting every vertex

Planar Graphs:

- Eulers formula: $v + f = e + 2$ if connected
- if atleast 3 sides to each face; $3f \leq 2e = \sum_{i=1}^{f} s_i$
- Not planar if $e > 3v - 6$
- Non-planar iff contains K_5 or $K_{3,3}$
- Handshake Lemma: sum of degrees $= 2|E|$

Complete Graphs

- contains maximum edges K_n
- K_n has $n(n-1)/2$ edges

biggest disconnected graph is where one vertex has no edge in a complete graph. $(n-2)(n-1)/2$

Trees

1) $n-1$ edges
2) removal of any edge disconnects graph
3) addition of any edge creates cycle.

Hypercubes

- vertex set of an n-dim hypercube is $V = \{0,1\}^n$ where its an n-bit string
- edge exists if two vertex differ by one bit string.
- 2^n vertices
- $n 2^{n-1}$ edges
- every vertex has degree n.
- view as bits

Midterm 2

Note 7: Public Key Cryptography

RSA:

1) Alice picks $N = p \cdot q$ where p, q large primes
2) Alice picks e relative prime to $(p-1)(q-1)$
   - public key (N, e)
3) Alice calculates private key $d = e^{-1} \bmod (p-1)(q-1)$
4) Bob takes message M and sends M^e.
5) Alice decrypts with d so $(M^e)^d = (M^{ed}) = M \bmod N$

Euler's Theorem: $a^{\varphi(n)} \equiv 1 \bmod n$ for n, a coprime

Fermats Little Theorem: $a^{p-1} \equiv 1 \bmod p$ for prime p and any int a.

Totient Properties: $- a^{(p-1)(q-1)} \equiv 1 \bmod pq$

$- \varphi(n) = n \prod_{i=1}^{k} \left(\frac{p_i - 1}{p_i}\right)$ $\quad - a^{p(p-1)} \equiv 1 \bmod p^2$

Note 8: Polynomials

- property 1: a non-zero polynomial of degree d has at most d roots
- property 2: given d+1 points, there's a unique polynomial of degree d (at most)
- degree d polynomial has d+1 coefficients so 10^{d+1} possible polynomials
- degree d polynomial mod m has m^{d+1} possibilities
- Polynomials of degree $\le d$ over F_m

# of points	# polynomials
$d+1$	1
d	m
$d-1$	m^2
$\vdots$	$\vdots$
$d-k$	m^{k+1}
$\vdots$	$\vdots$
0	m^{d+1}

Interpolation

- given d+1 points, $p(x) = \sum_{i=1}^{d+1} y_i \Delta_i(x)$ where $\Delta_i(x) = \frac{\prod_{j \ne i}(x - x_j)}{\prod_{j \ne i}(x_i - x_j)}$

Note 9: Error Correcting Codes

Erasure errors: send k extra points

General Corruption Errors: send 2k extra points

- Error-locator polynomial $E(x) = (x-e_1)(x-e_2)\cdots(x-e_k)$ for degree 1, $x + b_0$
- $Q(x) = a_{n+k-1}x^{n+k-1} + \cdots + a_1 x + a_0$

solve for system of equations: $Q(i) = r_i E(i)$ where r_i is the y value of each packet.

- find all coefficients a_i, the error is at position e_1, make sure to do $(-)$
- From $Q(x)$ and $E(x)$, compute $P(x) = \frac{Q(x)}{E(x)}$, and recompute the value at the error to get the correct message.

Note 10: Infinity and Computability

- one-to-one: if f maps to unique items
- onto: if f hits every element in the range
- bijection: 1-1 + onto
- set of all binary strings of any finite length $\{0,1\}^*$ is countable

Theorem

- $|P(N)| > |N|$

Note 11: Self Reference & Computability

Theorem: Halting Problem is uncomputable, ∃ doesn't program called TestHalt that outputs yes on input x on program P and otherwise No.

Easy Halting Problem: TestEasyHalt(P): yes if P halts on 0. no if P loops on 0.

- Doesn't exist because if it did we can make TestHalt(P, x)
  - construct a program P' that, on input 0, returns P(x), return TestEasyHalt(P')
- Disproving existence of program that determines whether program P on input x prints 'Hello World!'
  - Suppose it exists, then write program

```
reduce(input):
    execute(input)
    print 'Hello World!'
```

and run testHelloWorld(reduce(x))

Note 12: Counting

- when order matters, such as five card hand in poker, then $52 \times 51 \times 50 \times 49 \times 48$
- when order doesn't matter, just the set of fives, then $52C5 = \frac{52!}{47!5!}$
- $NCk = \frac{N!}{k!(N-k)!}$
- order matters and replacement with set $S = \{1, 2, \ldots, n\}$ and pick k elements: n^k
- order doesn't matter and replacement: $\binom{n+k-1}{k}$
  - then solve by looking at it as n bins, k elements
    - set up n-1 spaceholders, put k elements in it

Note 6 : Modular Arithmetic

Theorem 6.1: if $a \equiv c \mod m$ and $b \equiv d \mod m$ then $a+b \equiv c+d \mod m$ and $a \cdot b \equiv c \cdot d \mod m$

Theorem 6.2: Inverse exists iff $\gcd(m,x) = 1$. and inverse is unique

Theorem 6.3: Let $x \geq y > 0$. $\gcd(x,y) = \gcd(y, x \bmod y)$

Finding inverse through Euclid's Extended Algorithm

- inverse of $13x = 5 \mod 46$.

$46 = 13 \times 3 + 7$

$13 = 7 \times 1 + 6$

$7 = 6 \times 1 + 1$

$1 = 7 - 6$

start substituting

$1 = 7 - 13 - (7 \times 1)$

$1 = 2 \times 7 - 13$

$1 = 2 \times (46 - 13 \times 3) - 13$

$= 2 \times 46 - 7 \times 13$

$\therefore$ inverse is -7 so $-7 \bmod 46 \equiv 39$

- if there is atleast 1 solution to $ax = b \mod m$ where $d = \gcd(a,m)$, there are d solutions, each of form $x + i\frac{m}{d} \mod m$, distinct until $i\frac{m}{d} = m$ or when $i = d$

Note 7 : Public Key Cryptography

Fermats Little Theorem: For $\forall p \in$ prime, and any $a \in \{1, 2, \ldots, p-1\}$

$a^{p-1} \equiv 1 \mod p$.

Theorem 7.3 [Prime Number Theorem]: let $\pi(n)$ denote # primes that are less than or equal to n. Then for $n \geq 17$, $\pi(n) \geq n/\ln(n)$ and

$$\lim_{n \to \infty} \frac{\pi(n)}{n/\ln(n)} = 1$$

Conditional stuff for two R.V

- $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

and

- $P(a \leq X \leq b | Y = y) = \int_a^b f_{X,Y}(x|y)\,dx$

so - $E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$

$E[Y] = \int_{-\infty}^{\infty} E[Y|X=x] f_X(x)\,dx$

$E[X|B] = \sum_x x f(x|B)$

Arithmetic

Geometric: $\sum_{i=1}^{n} a_i = a\left(\frac{1-r^n}{1-r}\right)$

- Derangements $D_N = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right]$
- Probability that i passengers out of N sit in their assigned seats: (Derangement of N-i seats):

$\sum_{j=0}^{n-i} (-1)^j \binom{n}{i+j} \frac{(n-(i+j))!}{n!}$ Inclusion-Exclusion

- Proving Memoryless Property for exponential RV: $P(X > s+t | X > t) = P(X > s)$
  - Rearrange to prove $P(X > s+t) = P(X > s)P(X > t)$ instead, use conditional probability to get to this and simply just do algebra,
- Suppose you have a joint density distribution, and they ask you for $E(X)$ or $E(Y)$, then you get $P[Y > y]$ then do $\int_0^1 Pr[Y > y]\,dy$
- When asked the probability of two differences with different densities, just multiply the two densities, and then do AREA.
- AREA is really helpful. Double Integral !!
- if dart lands at point $(\frac{1}{2}, \frac{1}{2})$ what is the probability that the player was X:
  - $Pr[X | X \in (x, x+\delta)] = \frac{Pr[X \in (x, x+\delta) \cap X]}{Pr[X \in (x, x+\delta) \cap X] + Pr[X \in (x, x+\delta) \cap \bar{X}]}$ plug in δ for the pdf

Note 13: Introduction to Discrete Probability

<u>sample space</u>: outcome of a random experiment, the set of all of them Ω

<u>sample point</u>: the outcome of a random experiment

<u>probability space</u>: is a sample space together with a probability of each sample point

- non-negative
- total one

<u>Balls and Bins</u>

ex) N unlabeled balls, k labeled bins: think of N+k-1 candies, using k-1 dividers for k kids

k Balls	N Bins	
Distinct	Distinct	N^k
Identical	Distinct	$\binom{n+k-1}{n-1}$ or $\binom{n+k-1}{k}$
Distinct	Identical	$\sum_{i=1}^{n} S(k,i)$ where $S(k,n) = \frac{1}{n!}\sum_{i=0}^{n-1}(-1)^i \, {}_nC_i \, (n-i)^k$
Identical	Identical	$\sum_{i=1}^{n} P(k,i)$ [the # of partitions of int k into n parts

Note 14: Conditional Probability, Independence, and Combination of Events

- $P[A|B] = \frac{P[A\cap B]}{P[B]} = P$
- $P(A\cap B) = P(A|B)P(B) = P(B|A)P(A)$

general case

<u>Total Probability Rule for any event B</u>

- $P(B) = \sum_{i=1}^{n} P(B\cap A_i) = \sum_{i=1}^{n} P(B|A_i)P[A_i]$

<u>Bayes Rule</u>

- $P[A_i|B] = \frac{P[B|A_i]P[A_i]}{P[B]} = \frac{P[B|A_i]P[A_i]}{\sum_{j=1}^{n} P(B|A_j)P(A_j)}$

<u>Independence</u>: independent if $P[A\cap B] = P(A)P(B)$ or $P[A|B] = P(A)$

<u>Mutual Independence</u>: if every subset is independent.

<u>Product Rule</u>

- $P(A\cap B) = P(A)P(B|A)$
- $P(\cap_{i=1}^{n} A_i) = P[A_1] \times P[A_2|A_1] \times P[A_3|A_1\cap A_2] \times \cdots \times P[A_n | \cap_{i=1}^{n-1} A_i]$

<u>Inclusion-Exclusion</u>: let $A_1, \dots, A_n$ be events in some prob. space, then

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i\cap A_j) + \sum_{i<j<k} P(A_i\cap A_j\cap A_k) - \dots + (-1)^{n-1} P[A_1\cap A_2\cap \dots A_n]$$

<u>Mutual Exclusive Events</u>: if $A_1, \dots A_n$ are mutually exclusive $(A_i \cap A_j = \emptyset)$

then $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

<u>Union Bound</u>

$$P(\cup_{i=1}^{n} A_i) \le \sum_{i=1}^{n} P[A_i]$$

<u>Bayes Rule</u>

$$P[A|B] = \frac{P(A\cap B)}{P(B)} = \frac{P[B|A]P(A)}{P[B]}$$

<u>Total Probability Rule</u>

$$P[B] = P[A\cap B] + P[\bar{A}\cap B] = P[B|A]P(A) + P[B|\bar{A}](1-P(A))$$

- $Pr[C|B] = Pr[C|A\cap B]Pr[A|B] + Pr[C|\bar{A}\cap B]Pr[\bar{A}|B]$
- if A, B disjoint, then $P[A\cap B] = 0$
- if $P(A), P(B) > 0$ and A, B disjoint then dependent
- disjoint means $P(A|B) = 0$ so $P[A|B] = P(A)$ but can't be 0.
- $Pr[\bar{B}\cap\bar{G}] = 1 - Pr(B\cup G)$

Note 15: Random Variables: Distribution and Expectation

<u>Random Variable</u>: r.v X on sample space Ω is a function $X: \Omega \to \mathbb{R}$ that assigns to each sample point $\omega \in \Omega$ a real number $X(\omega)$

<u>Distribution</u>: distribution of a discrete random variable X is the collection of values $\{[a, P(X=a)]\} : a \in A$ where A is the set of all possible values of taken by X.

<u>Bernoulli Distribution</u>

$$P(X=i) = \begin{cases} p & \text{if } i=1 \\ 1-p & \text{if } i=0 \end{cases} \leftrightarrow X \sim \text{Bernoulli}(p)$$

<u>Binomial Distribution</u>

- X = # heads in n coin tosses, each H is probability p. $\leftrightarrow X \sim Bin(n,p)$

$$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i} \text{ for } i = 0, \dots, n$$

<u>Joint Distribution for variables X and Y</u>:

- $\{((a,b), P[X=a, Y=b]) : a\in A, b\in B\}$ where A is possible values of X and B p.v of Y.
- when given a joint distribution, $P[X=a]$ for X is the <u>Marginal Distribution</u>:

$$P(X=a) = \sum_{b\in A} P[X=a, Y=b]$$

- <u>Independence of r.v X and Y</u>:
  if $X=a$ and $Y=b$ are independent for values a, b.
  so $P(X=a, Y=b) = P[X=a]P[Y=b] \ \forall a, b$
- <u>Expectation</u> of a discrete random variable X:

$$E(X) = \sum_{a\in A} a \times P[X=a]$$

- $P[X=x, Y=y] = P[X=x|Y=y]P[Y=y]$

<u>Test Stuff</u>

- $R\cup(S\cap T) = (R\cup S)\cap(R\cup T)$
- $R\cap(S\cup T) = (R\cap S)\cup(R\cap T)$
- A, B independent $\to$ $\bar{A}$ and B independent and $\bar{A}, \bar{B}$ ind.
  - $P[\bar{A}|B] = 1 - Pr[A|B] = 1 - P(A) = Pr[\bar{A}]$
- $\gcd(x,y) = \gcd(y, x \bmod y)$ where $x > y$

<u>CRT</u>

1. compute $N = n_1 \times n_2 \times \dots \times n_k$
2. For each i, compute $y_i = \frac{N}{n_i}$
3. For each i, compute $z_i = y_i^{-1} \bmod n_i$
4. calculate $\sum_{i=1}^{k} a_i y_i z_i$

<u>Extended Euclid</u>

inverse of $13x = 5 \bmod 46$

$46 = 13\times 3 + 7$

$13 = 7\times 1 + 6$

$7 = 6\times 1 + 1$

$1 = 7 - 6$

then start substituting:

$1 = 7 - 13 - (7\times 1)$

$1 = 2\times 7 - 13$

$1 = 2\times(46 - 13\times 3) - 13$

$1 = 2\times 46 - 7\times 13$

Inverse is -7 so $39 \bmod 46$

<u>Packet Loss</u>

- Suppose we lose fraction f of packets. Then we send m packets and lose mf. so we must send $m = n + mf$

$m = n/1-f$

Post Midterm 2

Note 16: Random Variables: Variance and Covariance

Variance: $Var(X) = E[(X-\mu)^2] = E[X^2] - \mu^2$ where $E(X) = \mu$

- $var(cX) = c^2 var(X)$
- if $X_n = I_1 + I_2 + \dots + I_N$, then $E(X_n^2) = \sum_{i=1}^{n} E[I_i^2] + 2\sum_{i<j} E[I_i I_j]$
$= nE[I_i^2] + 2n(n-1)E[I_i I_j]$

Independent Random Variables
- $E(XY) = E(X)E(Y)$
- $var(X+Y) = var(X) + var(Y)$
- $cov(X,Y) = 0$ [converse is false]

Covariance: $cov(X,Y) = E(XY) - E(X)E(Y)$
- if $X \perp\!\!\!\perp Y$, $cov(X,Y) = 0$
- $cov(X,X) = var(X)$
- $var(X+Y) = var(X) + var(Y) + 2cov(X,Y)$
- covariance is bilinear: $cov(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \, cov(X_i, Y_j)$

Correlation: if $\sigma(X)$ and $\sigma(Y) > 0$, then
$corr(X,Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)}$

theorem: if $\sigma(X)$ & $\sigma(Y) > 0$, $-1 \le corr(X,Y) \le 1$

Note 18: Concentration Inequalities and the Law of Large Numbers

Markov's Inequality: For a nonnegative R.V X ($X(\omega) \ge 0 \ \forall \omega \in \Omega$) with finite mean,
- $P[X \ge c] \le E(X)/c$ for any positive constant c.

Proof: Let $\mathcal{l}$ denote the range of X and consider any constant $c \in \mathcal{l}$. Then,
$E(X) = \sum_{a \in \mathcal{l}} a \times P[X=a] \ge \sum_{a \ge c} a \times P[X=a] \ge \sum_{a \ge c} c \times P[X=a] = c\sum_{a \ge c} P[X=a] = cP[X \ge c]$

Generalized Markov's: Let Y be an R.V with finite mean. Then for constants c and $r \ge 0$,
- $P[|Y| \ge c] \le \frac{E[|Y|^r]}{c^r}$

Chebyshev's Inequality: for R.V X with $E[X] = \mu$,
- $P[|X - \mu| \ge c] \le \frac{Var(X)}{c^2}$ for any positive constant c.
- corollary: let $\sigma = \sqrt{var(X)}$ then [standard deviation]
  - $P[|X - \mu| \ge k\sigma] \le 1/k^2$
- confidence intervals: $P[|\hat{p} - p| \ge \varepsilon] \le \frac{Var(\hat{p})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$
  - we need $n \ge \frac{\sigma^2}{\varepsilon^2 \delta}$ where n is sample size, ε is error, δ is confidence.

Law of Large Numbers: let $X_1, X_2, \dots$ be a sequence of i.i.d R.V with common finite expectation $E[X_i] = \mu$ for all i. Then their partial sums $S_n = X_1 + X_2 + \dots + X_n$ satisfies
$P[|\frac{1}{n}S_n - \mu| < \varepsilon] \to 1$ as $n \to \infty$

for $\forall \varepsilon > 0$.

Note 19: Distributions (All)

1) Bernoulli Distribution: R.V that takes value 0 or 1

$P(X=i) = \begin{cases} p & \text{if } i=1 \\ 1-p & \text{if } i=0 \end{cases}$ where $0 \le p \le 1$

$X \sim Bernoulli(p)$

- $E(X) = p$ - $Var(X) = p(1-p)$

2) Binomial Distribution

$X \sim BIN(n,p)$ where n = # trials and p = probability of success for each trial

$P[X=i] = \binom{n}{i} p^i (1-p)^{n-i}$ for $i = 0,1,\dots,n$

- $E(X) = np$
- $Var(X) = np(1-p)$

3) Geometric Distribution

$X \sim Geometric(p)$ where p is the probability of success
- Models throwing a biased coin with Heads probability p X times until the first head appears

$P[X=i] = (1-p)^{i-1} p$ for $i = 1,2,\dots$

$P[X > k] = (1-p)^k$

- $E(X) = 1/p$
- $Var(X) = 1-p/p^2$
- Coupon Collector's Problem

$p_i = n-i+1/n$ and hence $E(X_i) = n/n-i+1$

$\therefore E(S_n) = \sum_{i=1}^{n} E[X_i] = \frac{n}{n} + \frac{n}{n-1} + \dots + \frac{n}{2} + \frac{n}{1} = n\sum_{i=1}^{n} 1/i$

where p_i is the probability of getting the i-th new coupon

Tail Sum Formula: Let X be a R.V that takes values $\{0,1,2,\dots\}$. Then $E[X] = \sum_{i=1}^{\infty} P[X \ge i]$

4) Poisson Distribution: R.V X for which

$P[X=i] = \frac{\lambda^i}{i!} e^{-\lambda}$ for $i = 0,1,2,\dots$

$X \sim Poisson(\lambda)$ where λ is the rate of an event occuring

- $E(X) = \lambda$
- $Var(X) = \lambda$

Independent Poissons: if $X \sim Poisson(\lambda)$ and $Y \sim Poisson(\mu)$ are independent, then $X+Y \sim Poisson(\lambda+\mu)$

Poisson as a limit of Binomial: Let $X \sim Binomial(n, \lambda/n)$ where $\lambda > 0$, Then $\forall i = 0,1,\dots,$ $P[X=i] \to \frac{\lambda^i}{i!}e^{-\lambda}$ as $n \to \infty$

General Tail Sum Formula: let R.V X be non-negative, then

$E[X] = \sum_{i=1}^{\infty} P[X \ge i]$ or

$E[X] = \int_0^{\infty} P(X > x)\,dx$

Note 20: Continuous Probability Distributions (R.V $\in \mathbb{R}$)

Probability Density Function: p.d.f for a realvalued R.V X is a function $\mathbb{R} \to \mathbb{R}$ satisfying
1) f is non-negative: $f(x) \ge 0 \ \forall x \in \mathbb{R}$
2) The total integral of f = 1: $\int_{-\infty}^{\infty} f(x)\,dx = 1$

- the distribution of X is given by $P(a \le X \le b) = \int_a^b f(x)\,dx \ \forall a < b$ (intervals)

Cumulative Distribution Function: c.d.f is a function $F(x) = P(X \le x) = \int_{-\infty}^{x} f(z)\,dz$
- $\frac{d}{dx}$ c.d.f = p.d.f : let c.d.f = $F(x)$ then p.d.f = $f(x) = \frac{d}{dx}F(x)$

Expectation: the expectation of a continuous R.V with p.d.f f is

$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$

Variance: the variance of a continuous R.V with p.d.f f is

$Var(X) = E[X^2] - E(X)^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \left[\int_{-\infty}^{\infty} x f(x)\,dx\right]^2$

Joint Density: A joint density function for 2 R.V X and Y is a function $f: \mathbb{R}^2 \to \mathbb{R}$ satisfying

1. f is nonnegative: $f(x,y) \geq 0 \ \forall x,y \in \mathbb{R}$
2. The total integral of $f = 1$: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

Joint Distribution of X & Y is given by:

$$P[a \leq X \leq b, c \leq Y \leq d] = \int_c^d \int_a^b f(x,y)\,dx\,dy \quad \forall a \leq b, c \leq d$$

"probability per unit area"

Independence for continuous R.Vs: Two continuous R.V with events $a \leq X \leq b$ and $c \leq Y \leq d$ are independent $\forall a \leq b, c \leq d$ if

- $P[a \leq X \leq b, c \leq Y \leq d] = P[a \leq X \leq b] \cdot P[c \leq Y \leq d]$

Theorem: if $X \perp\!\!\!\perp Y$, $f(x,y) = f_X(x) f_Y(y) \ \forall x,y \in \mathbb{R}$

(Joint Density of $\perp\!\!\!\perp$ R.V's $= \prod$ marginals densities)

5) Exponential Distribution

For $\lambda > 0$, a continuous R.V X with p.d.f $f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$ is called an

$X \sim Exp(\lambda)$: continuous version of geometric distribution

- (λ is the rate at which an event happens)
- (how long it takes something to happen)
- (rate) - (Time)

- $E(X) = 1/\lambda$
- $Var(X) = 1/\lambda^2$
- For any $t \geq 0$, $P[X \geq t] = P(X > t] = \int_t^{\infty} \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\Big|_t^{\infty} = e^{-\lambda t}$

(probability that we have to wait more than time t for our event to happen is $e^{-\lambda t}$, exponential decay with rate λ)

Poisson Arrival Process: For $i = 1, 2, \ldots$, let W_i denote the waiting time to the ith arrival, Then

1. $W_i \sim Exp(\lambda) \ \forall i = 1, 2, \ldots$
2. $W_1, W_2, \ldots$, are mutually independent

6) Normal Distribution

For any $\mu \in \mathbb{R}$ and $\sigma > 0$, a continuous R.V X with p.d.f

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

μ = mean
σ^2 = variance

is called a normal random variable with parameters μ and σ^2

$X \sim N(\mu, \sigma^2)$. in the case $\mu = 0$ and $\sigma = 1$, X is the standard Normal distribution

Lemma: if $X \sim N(\mu, \sigma^2)$ then $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$
and if $Y \sim N(0,1)$ then $X = \sigma Y + \mu \sim N(\mu, \sigma^2)$

- $E(X) = \mu$
- $Var(X) = \sigma^2$

Sum of independent Normal variables: let $X \sim N(0,1)$ and $Y \sim N(0,1)$ be independent standard normal random variables, and suppose $a, b \in \mathbb{R}$ are constants. Then $Z = aX + bY \sim N(0, a^2 + b^2)$

Corollary: let $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ be independent normal random variables. Then for any constants $a, b \in \mathbb{R}$, the R.V $Z = aX + bY$ is also normally distributed with mean $\mu = a\mu_X + b\mu_Y$ and variance $\sigma^2 = a^2\sigma_X^2 + b^2\sigma_Y^2$

Central Limit Theorem: let $X_1, X_2, \ldots$ be a sequence of i.i.d R.V with common finite expectation $E(X_i) = \mu$ and finite variance $Var(X_i) = \sigma^2$. Let $S_n = \sum_{i=1}^n X_i$, Then, the distribution of $S_n - n\mu / \sigma\sqrt{n}$ converges to $N(0,1)$ as $n \to \infty$,

$$P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq c\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} e^{-x^2/2}\,dx \quad \text{as } n \to \infty$$

Buffon's Needle Probability: $2/\pi$

- can proving using integration of JDF: $\int_{-\pi/2}^{\pi/2} \int_0^{\ell/2} \frac{2}{\pi\ell}\,dy\,d\theta$
- can prove using indicator var's for circle

Note 21: Finite Markov Chains

* Can also use Markov Chain First Step for probability. End state $= 1$

General Finite Markov Chain

- state space: $\mathcal{S} = \{1, 2, \ldots, k\}$ for some finite k
- transition probability matrix $P = |\mathcal{S}| \times |\mathcal{S}|$
- initial distribution row vector μ^0

Theorem: for all $n \geq 0$, $\mu^{(n)} = \mu^{(0)} P^n$ in particular, if $\mu_i^{(0)} = 1$, for some i, then $\mu_j^{(n)} = [P^n]_{ij} = P[X_n = j \mid X_0 = i]$

First Step Analysis: solve a system of equations by solving for all $\tau(A), \ldots, \tau(i)$'s.
Let $\tau(i)$ be the average number of steps until the Markov Chain enters one of the states in $\mathcal{A}$, given it starts at i.
Then,
$\tau(i) = 0$ if $i \in \mathcal{A}$
$\tau(i) = 1 + \sum_{j \in \mathcal{S}} P_{ij} \tau(j)$ where $\mathcal{A} \subset \mathcal{S}$ is a subset of states

Stationary or Invariant Distribution: A distribution $\pi = (\pi_i : i \in \mathcal{S})$ is invariant or stationary for the transition prob matrix P if it satisfies $\pi = \pi P$.

Theorem: the distribution $\mu^{(n)} = \mu^{(0)} P^n$ satisfies $\mu^{(n)} = \mu^{(0)}$ for all $n \in \mathbb{N}$ if $\mu^{(0)}$ is invariant

Irreducible: irreducible M.C if it can go from every state $i \in \mathcal{S}$ to every other state $j \in \mathcal{S}$ possibly in multiple steps

Theorem: if a MC with a finite state space $\mathcal{S}$ and transition probability matrix P is irreducible, then for any initial distribution $\mu^{(0)}$ and for all $i \in \mathcal{S}$

$$\frac{1}{n}\sum_{m=0}^{n-1} I\{X_m = i\} \to \pi_i \quad \text{as } n \to \infty$$

where $\pi = (\pi_i : i \in \mathcal{S})$ is an invariant distribution for P.
Consequently, the invariant distribution exists and is unique

Theorem: consider an irreducible MC on $\mathcal{S}$ with transition prob matrix P. For $i \in \mathcal{S}$ define $d(i) = gcd\{n > 0 \mid [P^n]_{ii} = P[X_n = i \mid X_0 = i] > 0\}$

1) Then, $d(i)$ has the same value for all $i \in \mathcal{S}$. if that value is 1, then MC is aperiodic. Otherwise, periodic with period d.
2) iff an irreducible MC is aperiodic, then $\forall i \in \mathcal{S}$,
$P[X_n = i] \to \pi_i$ as $n \to \infty$
where $\pi = (\pi_i : i \in \mathcal{S})$ is unique invariant distribution for P.

7) Uniform Distribution

Distribution that represents an event that randomly happens at any time during an interval of time

- $f(x) = \frac{1}{b-a}$ for $a \leq x \leq b$ (density)
- $F(x) = 0$ for $x < a$, $\frac{x-a}{b-a}$ for $a < x < b$, 1 for $x > b$
- $E(x) = a+b/2$
- $Var(x) = \frac{1}{12}(b-a)^2$

$E(X^2) = 1/3$

- Joint pdf of two independent ones: $f_{A,B} = f_A \times f_B$

- showing 2 R.Vs are independent from Joint density: $f(x)$ must $= f(x \mid Y=y) = \frac{f(x,y)}{\int f(x,y)\,dx}$ ($\int f(x,y)\,dy$)
- law of iterated Expectation: $E[X] = E[E(X|Y)]$
- example of $cov(X,Y) = 0 \to$ not $\perp\!\!\!\perp$: $(X,Y) = \{(1,-1), (-1,1)\}$ $1/3$ each and $\{(\sqrt{2},\sqrt{2}), (-\sqrt{2},-\sqrt{2})\}$ $1/6$ each
- seq of R.V's are a M.C if $P[X_{t+1} \mid X_t] = P[X_{t+1} \mid X_t, X_{t-1} \cdots 0]$ $\forall t$ and $P[X_{t'+1} \mid X_{t'}] = P[X_{t+1} \mid X_t]$
- $E[X+Y \mid X+Y=1] = E[X \mid X+Y] + E[Y \mid X+Y] = 1.5$
- When they ask for pdf of min or max, or whatev, find $Pr[Y \geq y]$ then take derivative $(1-y)^n$; find $Pr[Y \leq y] = y^n$ = CDF
- ...
- using CLT, 95% confidence within 1 percent accuracy of p: $p \pm 2\sigma$ where $\sigma = \sqrt{\text{variance}}$. we wish $2\sigma \leq 0.01$, solve for N. σ should be in terms of N.
- Finding distribution of R.V $Z = X+Y$: First look at $P[Z=k]$
$P[Z=k] = \sum P(Z=k, Y=i) = \sum P(X=k-j, Y=i) = \sum (P(X=k-j))(P(Y=i))$
$= \frac{1}{N} \ldots$
- Getting cdf, by definition is getting $P[Z \leq z]$, but you can set $1 - P[Z \geq z]$ if easier
- Find joint density, then just double integral the area you don't want. Follow everything. When throwing multiple times MINTUR.
- uniform distribution = number line !!!
- $E[X|Y] = E[X \mid Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$