EECS 126 Probability and Random Processes — University of California, Berkeley: Fall 2014
Kannan Ramchandran — December 16, 2014

Final Exam

Last name	First name	SID

Rules.

- DO NOT open the exam until instructed to do so.
- Note that the test has 110 points. The maximum possible score is 100.
- You have 10 minutes to read this exam without writing anything.
- You have 150 minutes to work on the problems.
- Box your final answers.
- Partial credit will not be given to answers that have no proper reasoning.
- **Remember to write your name and SID on the top left corner of every sheet of paper.**
- **Do not write on the reverse sides of the pages.**
- All eletronic devices must be turned off. Textbooks, computers, calculators, etc. are prohibited.
- No form of collaboration between students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- **You must include explanations to receive credit.**

Problem	Part	Max	Points	Problem	Part	Max	Points
1	(a)	5		2		20	
	(b)	5		3		18	
	(c)	5		4		10	
	(d)	5		5		17	
	(e)	5		6		20	
		25					
Total						110	

Cheat sheet

1. Discrete Random Variables

1) Geometric with parameter $p \in [0,1]$:
   $P(X = n) = (1-p)^{n-1}p,\ n \geq 1$
   $E[X] = 1/p,\ \text{var}(X) = (1-p)p^{-2}$

2) Binomial with parameters N and p:
   $P(X = n) = \binom{N}{n}p^n(1-p)^{N-n},\ n = 0, \ldots, N,$ where $\binom{N}{n} = \frac{N!}{(N-n)!n!}$
   $E[X] = Np,\ \text{var}(X) = Np(1-p)$

3) Poission with parameter λ:
   $P(X = n) = \frac{\lambda^n}{n!}e^{-\lambda},\ n \geq 0$
   $E[X] = \lambda,\ \text{var}(X) = \lambda$

2. Continuous Random Variables

1) Uniformly distributed in $[a,b]$, for some $a < b$:
   $f_X(x) = \frac{1}{b-a}1\{a \leq x \leq b\}$
   $E[X] = \frac{a+b}{2},\ \text{var}(X) = \frac{(b-a)^2}{12}$

2) Exponentially distributed with rate $\lambda > 0$:
   $f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$
   $E[X] = \lambda^{-1},\ \text{var}(X) = \lambda^{-2}$

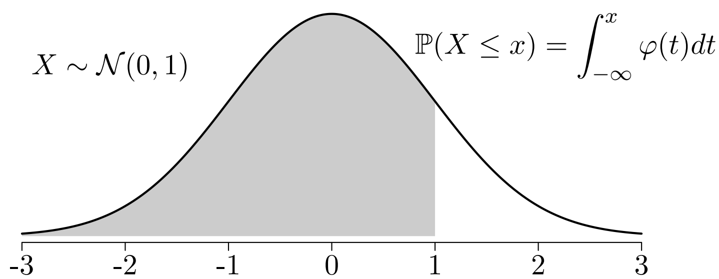
3) Gaussian, or normal, with mean μ and variance σ^2:
   $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$
   $E[X] = \mu,\ \text{var} = \sigma^2$

3. Estimation

1) LLSE: Let X and Y be random variables. Then,
   $L[X|Y] = E(X) + \frac{\text{cov}(X,Y)}{\text{var}(Y)}(Y - E[Y])$.

NAME: SID:

4. Normal Distribution Table



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Problem 1. Short questions.

(a) Given X, the random variables $Y_1, Y_2, \ldots, Y_n$ are i.i.d. and exponentially distributed with mean X. Suppose that $\Pr(X = 1) = 1 - \Pr(X = 2) = p \in (0, 1)$. Find $MAP[X|Y_1, Y_2, \ldots, Y_n]$.

(b) Consider two random variables X and Y. Is the following statement true or false.
If $L[X|Y] = E[X|Y]$, then X and Y are jointly Gaussian. Either argue that it is correct, or provide a counterexample.

(c) Find an example of two Gaussian random variables which are not jointly Gaussian.

(d) Consider a Poisson process $\{N_t, t \geq 0\}$. Let T_n be the random variable denoting the time of n-th arrival. Find $E[T_2|T_{10} = t]$.

(e) Consider a ternary tree as shown in Figure 1. Suppose that each edge in this tree is connected with probability p. We want to find the probability that the tree rooted at R is at least of depth n. Find a recursive formula to evaluate this probability.

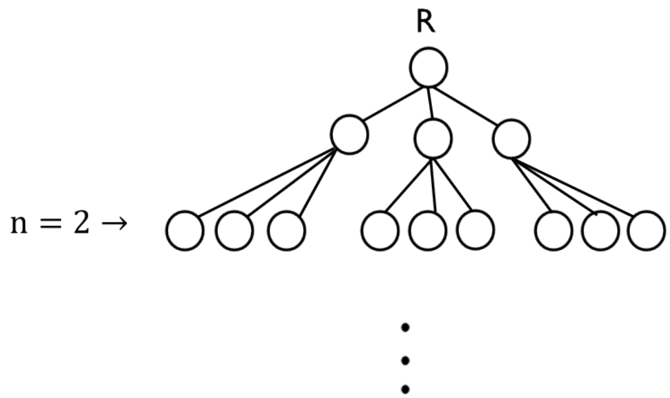


Figure 1: Ternary tree.

Problem 2. Alice shows up at an Internet cafe at time 0 and spends her time exclusively typing emails. The times that her emails are sent are modeled by a Poisson process with rate λ_A emails per hour.

(a) Let Y_1 and Y_2 be the times at which Alice's first and second emails are sent. Find $E[Y_2|Y_1]$ and the joint pdf of Y_1 and Y_2.

(b) You don't know λ_A but you watch Alice for an hour and observe that she has sent exactly 5 emails. Derive the maximum likelihood estimator of λ_A based on this.

(c) Let N be the number of emails sent during $[0,1]$. If N is modeled as normal random variable, and you observe the value of N, give an approximate 95% confidence interval for λ_A.

Bob shows up at time 1 and sits next to Alice. He starts typing emails at time 1, and sends them off according to an independent Poisson process with rate λ_B.

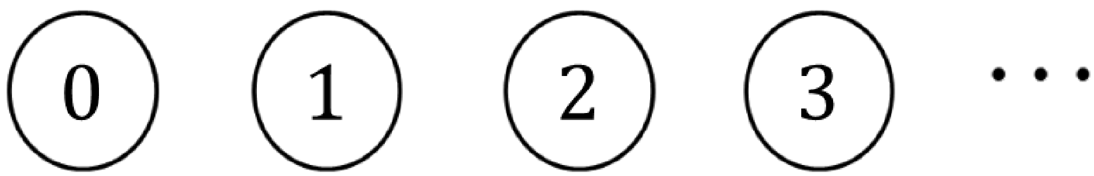
(d) What is the PMF of the total number of emails sent by Alice and Bob together during the interval $[0,3]$.

(e) What is the expected value of the total typing time associated with the email Alice is typing at the time Bob shows up? Note that "total typing time" includes the time that Alice spent on that email both before and after Bob's arrival.

(f) Given that a total of 10 emails were sent during the interval $[0, 2]$, what is the probability that exactly 4 of them were sent by Alice?

Problem 3. An online dating website tries to match couples. (We have a liberal attitude where couples can be of any gender.) Let X_n be the number of members of this site at time slot n. We want to analyze the discrete-time process $\{X_n, n \geq 0\}$. At each time slot, exactly one of the following events happens: (i) Two persons are happily matched and leave the website forever with probability p, (ii) A single frustrated person leaves the system individually with probability q, and (iii) a new member joins the system with probability $r = 1 - p - q$. If there is only one member in the system, that member leaves with probability $p + q = 1 - r$.

(a) Draw the transition diagram for this Markov chain, X_n.



(b) Write the balance equations for state 2 without solving them.

(c) Formally prove that the Markov chain is transient if $r > 2p+q$. [Hint: Try to characterize the random variable $X_n - X_{n-1}$, and use law of large numbers.]

Problem 4. Smart Alec thinks he has uncovered a good model for the relative change in daily stock price of XYZAir Inc., a publicly traded airline company in the New York Stock Exchange. His model is that the relative change in price, X, depends on the relative change in price of oil, Y, as well as some unpredictable factors, modeled collectively as a random variable Z. In his model, Y and Z are continuous iid random variables uniformly distributed between -1 and 1, and $X = Y + 2Z + Y^2$.

(a) Smart Alec first decides to use a Linear Least Square Estimator of X given Y. Find $L[X|Y]$. What is the MSE of Smart Alec's LLSE?

(b) Smart Alec now decides to use a more sophisticated quadratic least squares estimator for X given Y, i.e. an estimator of the form $Q[X|Y] = aY^2 + bY + c$. Find the set of equations that can be solved to find $Q[X|Y]$. Draw a geometric interpretation. Which estimator has a lower MSE?

Problem 5. Shown below is a state transition diagram corresponding to a rate- $1/2$ convolutional encoder having one delay. Recall from Lab 11 that this corresponds to an input message sequence $x[n]$, and two output sequences ($y_0[n], y_1[n]$). In the diagram, the state represents the value of $x[n-1]$ and the the transitions represent $x[n]\,/\,y_0[n]y_1[n]$. For example, "1/01" means " $x[n] = 1, y_0[n] = 0, y_1[n] = 1$ ".

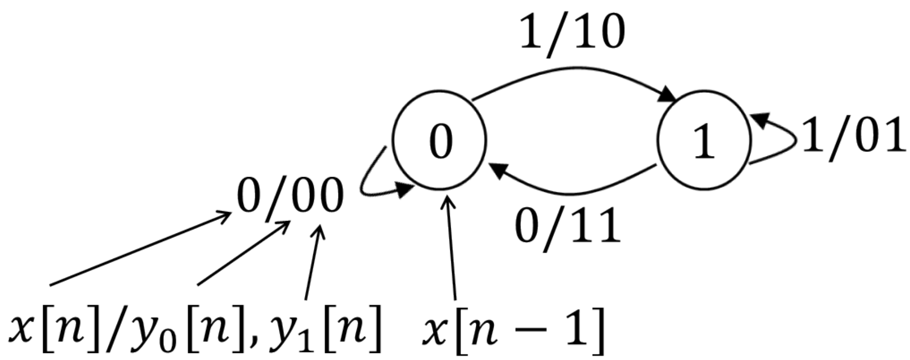


Figure 2: State transition diagram.

(a) Draw the simplest convolutional encoder corresponding to the given state transition diagram. Label the inputs and outputs clearly.

(b) Draw one stage of the trellis-diagram corresponding to the above state transition, where the branches correspond to the outputs emitted from each state.

(c) Assume you start at state 0, i.e. $x[0] = 0$. What is the encoded output $\{y_0[n], y_1[n]\}$ corresponding to the 4-bit input message sequence $x[n] = (1, 0, 1, 1)$ for $n = 1$ to $n = 4$?

(d) Again assume that $x[0] = 0$. A 4-bit message sequence $x[n]$ (for $n = 1$ to $n = 4$) is encoded, interleaved, and transmitted over a Binary Symmetric Channel with bit-flip probability equal to 0.1. See Figure 3. We receive the output sequence (10 11 01 10) corresponding to $r_0[n]$ $r_1[n]$ for $n = 1$ to $n = 4$. Assuming the message sequences transmitted have uniform priors, what is the MAP estimate of the transmitted 4-bit message sequence? (Clearly show your work to get credit).

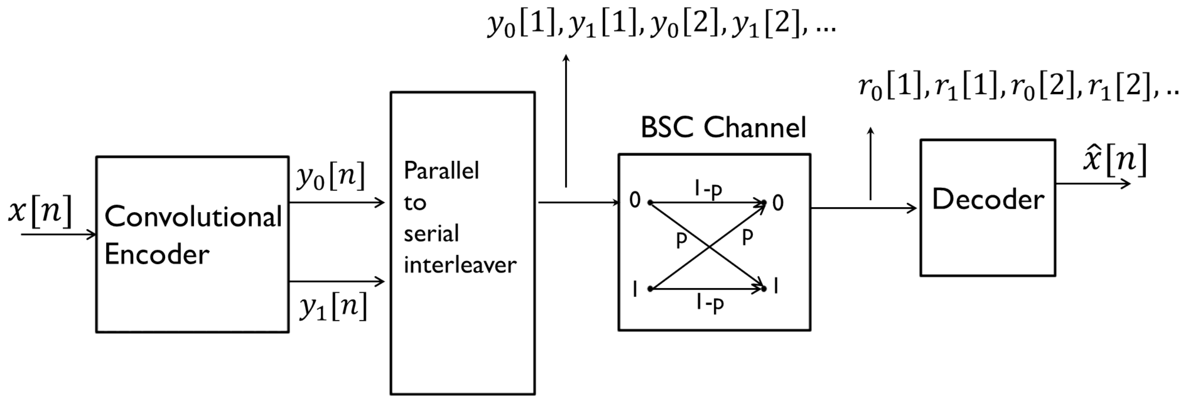


Figure 3: Convolutional Encoder.

Problem 6. There are two coins A and B having unknown biases θ_A and θ_B (these are the probabilities of being Heads for the two coins, respectively). Our goal is to estimate these two biases by repeating the following procedure 3 times:
1) Randomly (i.e. with probability 50%) choose one of the coins.
2) Perform 3 independent coin tosses with the chosen coin, and record the Head/Tail (H/T) sequence of tosses.

We will thus have 9 coin tosses in total, consisting of 3 sets of 3 tosses. Suppose your experiment has the following outcome: note that z_i denotes the label of the coin in set i, and y_i is the actual sequence of Head/Tail tosses in set i:

$$\begin{aligned} z_1 &= B, \quad y_1 = (H, T, T) \\ z_2 &= A, \quad y_2 = (H, H, H) \\ z_3 &= A, \quad y_3 = (H, T, H) \end{aligned}$$

(a) What is the maximum likelihood estimate of θ_A and θ_B if you are given access to both the coin label sequence z_1, z_2, z_3 and the coin toss outcome vectors y_1, y_2, y_3 in the above experimental observation?

(b) Now you are asked to find the maximum likelihood estimate of the bias parameters but given access only to the y_i information (in words, you do not know the labels of the coins that were tossed in each set, but know only the H/T toss sequence in each set). Explain briefly how you would find the (exact) maximum likelihood estimate if you have no constraints on computational complexity.

(c) In the interests of computational efficiency, you decide to use the EM algorithm to do your estimation. Suppose you start with initial estimates $\theta_A^0 = \frac{2}{3}$, and $\theta_B^0 = \frac{1}{3}$. What would one iteration of the EM algorithm produce as an estimate for θ_A^1 and θ_B^1 for the "hard" EM algorithm as well as the "soft" EM algorithm. Clearly explain your work to get credit.

[Hint: Recall that the E-step involves computing a probability distribution over all possible "completions" of the incomplete data using the current parameters, while the M-step involves determining new parameters using the current "completions."]

(d) Does the hard EM always converge to the MLE? Does the soft EM always converge to the MLE?

END OF THE EXAM.

Please check whether you have written your name and SID on every page.

Hope you enjoyed EE126 a lot; you learned a lot.