Assignments for IPL-10 (CS-723)

1 Probability Theory

1. Consider an imaginary situation where there is an infinite (2d) floor which is divided into square grids of side length l. Suppose we drop a needle of length $l/2$ on the floor. Making reasonable assumptions (say, "equally likely" etc.) arrive at the probability that the needle intersects a grid line. Clearly mention the sample space, event space and the required probability.

2. Let $\Omega = \mathbb{R}$ and $\mathcal{F}$ be the collection of all events of the form $(a, b] = \{x \in \mathbb{R} | a < x \leq b,\ a, b \in \mathbb{R}\}$ and their finite unions. By convention, assume that $(a, \infty) \in \mathcal{F}$ for $a \in \mathbb{R}$. Is $\mathcal{F}$ closed wrt. complementation? If so, is $\mathcal{F}$ a σ-algebra over $\mathbb{R}$? Justify.

3. Derive the following from the axioms of probability:

   (a) If $E_i \in \mathcal{F}, i = 1, \ldots, n$, then

      i. prove that:

$$P(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i) - \sum_{1 \leq i < j \leq n} P(E_i \cap E_j) + \sum_{1 \leq i < j < k \leq n} P(E_i \cap E_j \cap E_k)$$
$$+ \cdots + (-1)^n \sum_{i=1}^{n} P(\cap_{j=1, j \neq i}^{n} E_j) + (-1)^{n+1} P(\cap_{i=1}^{n} E_i)$$

      ii. prove the following union bounds on probability:

$$\sum_{i=1}^{n} P(E_i) - \sum_{1 \leq i < j \leq n} P(E_i \cap E_j) \leq \underbrace{P(\cup_{i=1}^{n} E_i) \leq \sum_{i=1}^{n} P(E_i)}_{Boole's\ inequality}$$

iii. prove the *Bonferroni's inequality*:

$$P(\cap_{i=1}^{n} E_i) \geq \sum_{i=1}^{n} P(E_i) - n + 1$$

4. Suppose $([0,1], \mathcal{F}, P)$ is a probability space where $\mathcal{F}$ is a σ-algebra over $[0,1]$. Consider a function $X : [0,1] \mapsto \mathbb{R}$ defined as $X(\omega) = \omega \ \forall \ \omega \in [0,1]$. Describe the minimal $\mathcal{F}$ such that X is a random variable. With this minimal $\mathcal{F}$ consider two random variables, X (defined above) and $Y : [0,1] \mapsto \mathbb{R}$ defined as $Y(\omega) = 1 - \omega \ \forall \ \omega \in [0,1]$ (Is Y a random variable ?). Plot the distribution functions of X, Y.

5. Show that a discrete random variable X taking on values from $\mathbb{N}$ (set of Natural numbers) can satisfy the following property (known as "memory-less" property): $P[X > m+n/X > n] = P[X > m] \ m, n \in \mathbb{N}$ if and only if it follows a geometric distribution.

6. Consider n IIB trials of tossing a coin where probability of getting a head is $p \in (0,1)$. Let the "big" probability space describing these n expts. with probability function defined "consistently" be $\mathbb{P}_{IIB}^{n}$[1]. Consider two random variables defined on $\mathbb{P}_{IIB}^{n}$: $X =$ the number of times two consecutive heads occurred; $Y =$ the number of tosses required to see two consecutive heads for the first time (note that X, Y are generalizations of the binomial and geometric rvs). Attempt writing down the pmfs of X, Y.

7. Consider infinite tosses of coins (i.e., $\mathbb{P}_{IIB}^{\infty}$). What is the probability that the pattern THHH occurs before HHHH ?

8. Let A and B be any two subsets of $S = \{1, \ldots, n\}$ picked up at "random". What is probability of the event $A \subset B$ and the event $A \cap B = \phi$?

[Due: 16-Aug]

9. Let X follow Normal distribution. Define a new rv: $Y = X^2$. Compute the pdf of Y (the name given to the distribution of Y is "chi-square"). Now define another rv: $Z = \sqrt{Y}$. Compute the pdf of Z. (Note that distributions of Z and X are different; infact the distribution of Z is known as "chi".).

[1] This notation will be used through-out this text.

10. Suppose a projectile is fired with a fixed velocity v and at a random angle Θ which follows a uniform distribution between $[0, \frac{\pi}{2}]$. Let R represent the random variable: "range of the projectile". What is the distribution of R (specify its pdf) ?

11. Compute all moments and absolute moments of a std. Normal rv.

12. Show that median is $\text{argmin}_c \mathbb{E}[|X - c|]$.

13. Let X, Y be two continuous rvs and f_X, f_Y denote their pdfs. Show that $\mathbb{E}[\log(f_X(X))] \geq \mathbb{E}[\log(f_Y(X))]$.

14. Suppose an urn has $m = m_1 + m_2$ balls where m_1 are blue and m_2 are red. Consider the random expt., where balls are picked (independently with replacement) until, a ball of complementary color as that of the first draw is picked. Consider the rv X which is the number of balls picked. Compute $\mathbb{E}[X]$. Can you recompute $\mathbb{E}[X]$ in the case without replacement ?

15. If X is a std. Normal rv. show that $P[X \geq \epsilon] \leq \exp(-\epsilon^2/2)$. (Hint: Think in terms of mgf)

[Due: 03-Sep]

16. Consider a function of two rvs X, Y given by: $Z = \frac{\min(X,Y)}{\max(X,Y)}$. Suppose X, Y take only non-negative values. Derive an expression for pdf of Z in terms of joint pdf of X, Y. Simplify this expression for the case X, Y are iid rvs with exponential distribution.

17. Show that: $var(X) = \mathbb{E}[var(X/Y)] + var(\mathbb{E}[X/Y])$.

18. Given $\{X_n\} \xrightarrow{r} X$ show that $\{X_n\} \xrightarrow{s} X \; \forall \; 0 < s < r, \; s \in \mathbb{N}$.

19. Let $\{X_n\}$ be a sequence of iid random variables following the (discrete) uniform distribution. The discrete set of values each of the rvs X_n take is $\{0, 1, \ldots, 9\}$. Now, consider a new sequence of rvs defined as: $U_n = \sum_{i=1}^{n} \frac{X_i}{10^i}$. Prove that (being as rigorous as possible) $\{U_n\} \xrightarrow{\mathcal{D}} U$, where U is a (continuous) uniform random variable between 0 and 1[2].

[2] This is an eg. of case where a sequence of discrete rvs converges to a continuous rv! Also, realize that U_n is simply decimal expansion (upto n decimals) of a number between 0 and 1

20. Find probability that a fair coin lands heads 20 times when flipped for 40 times using: a) Normal approx. (CLT) b) Exact soln. How close are these values ? (Assume $\Phi(0.16) = 0.5636$).

21. The weights of a population of workers have (true) mean 167 and (true) std.div. 27. If a sample of 36 workers is chosen, what is (approx.) the probability that the sample mean (avg. values of the samples) of their weights is between 163 and 170 ? Repeat the question when sample size is 144 workers. (Assume $\Phi(0.6667) = 0.7475, \Phi(-0.8889) = 0.1870, \Phi(1.3333) = 0.9088, \Phi(-1.7778) = 0.0377$).

22. X_1, X_2, X_3 are three independent std. Normal rvs. Let $Y_1 = 3X_1 + 5X_2 + X_3, Y_2 = 2X_1 + 5X_3$. Write down the joint-pdf of Y_1, Y_2, if it exists.

23. Consider infinite sequence of iid rvs $X_1, \ldots, X_n, \ldots$. Consider the following rv $S_n = (\sum_{i=1}^{n}(X_i - \frac{\sum_{j=1}^{n} X_j}{n})^2)/n$ (which is the usual way of computing sample variance of n numbers). Is S_n an estimator of $var(X_i)$? If so is it unbiased ?

24. Consider a multivariate rv $X = [X_1 \ldots X_n]^\top$ for which the mean vector is zero. Compute (in terms of n) $\mathbb{E}[X\Sigma^{-1}X^\top]$ where Σ is the covariance matrix (invertible) of X.

[Due: 12-Oct]