ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 1 | Information Theory and Coding
Homework 1 | September 21, 2010, SG1 – 15:15-17:00

Problem 1 (Series)**.** Compute the following sums and list the conditions for convergence: (a) $\sum_{i=0}^{n} \alpha^i$ (b) $\sum_{i=0}^{\infty} \alpha^i$ (c) $\sum_{i=1}^{\infty} i\alpha^i$. Here α is a complex number and n is a non-negative integer.

Problem 2 (Bayes' Theorem)**.** Only three companies A, B, and C manufacture a certain product. Company A has a 50% market share and company B has 30% market share. 5% of company A's products are defective, 7% of company B's products are defective, and 10% of company C's products are defective.

(a) What is the probability that a randomly selected product is defective?

(b) What is the probability that a defective product came from company B?

(c) Is the event that a randomly selected product is defective independent from the event that a randomly selected product came from company B?

Problem 3 (Probability Distributions)**.**

(a) (Geometric Distribution) If X is a discrete random variable with geometric distribution, then $\Pr\{X = t\} = p(1-p)^t$, $t = 0, 1, 2, \cdots$, where p is a given parameter of the distribution such that $0 < p < 1$. An example of such a distribution is when X is the number of tails before heads comes up in repeated tossing of a biased coin with probability of heads equal to p.

  (i) Find $E[X]$ and $Var[X]$.

  (ii) Find the cumulative distribution function $\Pr\{X \le t\}$.

  (iii) (Memoryless Property) A random variable X is memoryless with respect to $t > 0$, if for any $s > 0$, $\Pr\{X > s + t | X \ge t\} = \Pr\{X > s\}$. Show that, if s and t are non-negative integers, the geometric distribution satisfies the memoryless property. [Hint: Show that $\Pr\{X > s + t\} = \Pr\{X > s\} \Pr\{X \ge t\}$. Then use the fact that $\Pr\{X > s + t\} = \Pr\{X > s + t, X \ge t\}$.]

  (iv) If X is a memoryless random variable, show that $E[X - t | X \ge t] = E[X]$.

(b) (Binomial Distribution) If X is a discrete random variable with binomial distribution, then $\Pr\{X = n|N\} = \binom{N}{n} p^n (1-p)^{N-n}$, $n = 0, 1, \cdots, N$, where N, a non-negative integer, and p, $0 < p < 1$, are given parameters of the distribution. An example of such a distribution is when X is the number of heads in N repeated tosses of a biased coin with probability of heads equal to p. Find $E[X]$ and $Var[X]$.

(c) (Poisson Distribution) If X is a non-negative discrete random variable with Poisson distribution, then $\Pr\{X = n\} = e^{-\lambda} \frac{\lambda^n}{n!}, n = 0, 1, 2, \cdots$, where $\lambda > 0$ is a given parameter of the distribution.

  (i) Find $E[X]$ and $Var[X]$.

(ii) Setting $\lambda = Np$, show that the Poisson distribution is a limiting case of the binomial distribution, that is,

$$\Pr\{X = n\} = \lim_{N\to\infty} \Pr\{X = n|N\}.$$

(d) (Exponential Distribution) If X is a non-negative continuous random variable with exponential distribution, then the probability density function of X is given by $p_X(x) = \lambda e^{-\lambda x}$, $x \geq 0$, where $\lambda > 0$ is a given parameter of the distribution.

(i) Find $E[X]$ and $Var[X]$.

(ii) (Memoryless Property) Show that the exponential distribution satisfies memoryless property, that is, if X is a random variable with exponential distribution, then for any $s, t > 0$, $\Pr\{X > s + t|X \geq t\} = \Pr\{X > s\}$.

(iii) If X and Y are independent random variables having exponential distributions with parameters λ_X and λ_Y respectively, then show that the random variable $Z = \min\{X, Y\}$ has an exponential distribution with parameter $\lambda_X + \lambda_Y$.

(e) (Gaussian Distribution) If X is a continuous random variable with Gaussian distribution, then the probability density function of X is given by $p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$, where μ and σ are given parameters are of the distribution. It is denoted as $X \sim \mathcal{N}(\mu, \sigma^2)$. Find $E[X]$ and $Var[X]$.

(f) (Multivariate Gaussian) Let $\mathbf{X} = (X_1, X_2, \cdots, X_n)^T$ be a vector of n random variables. It is said to have a multivariate normal distribution, denoted as $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$, if any of the following *equivalent* conditions hold:

I. There exists an n-dimensional vector μ and a positive semi-definite $n \times n$ matrix Σ such that for any n-dimensional vector $\mathbf{a}$, the linear combination $\tilde{X} = \mathbf{a}^T\mathbf{X}$ has the following univariate Gaussian distribution: $\tilde{X} \sim \mathcal{N}(\mathbf{a}^T\mu, \mathbf{a}^T\Sigma\mathbf{a})$.

II. There exists an n-dimensional vector μ and a positive semi-definite $n \times n$ matrix Σ such that $\mathbf{X}$ has the characteristic function (that is, the Fourier transform of its probability density function $p_\mathbf{X}(\mathbf{x})$) given by

$$\phi_\mathbf{X}(\mathbf{t}) = E[e^{i\mathbf{t}^T\mathbf{X}}] = e^{i\mathbf{t}^T\mu - \mathbf{t}^T\Sigma\mathbf{t}}.$$

If Σ is positive definite, then its probability density function is given by

$$p_\mathbf{X}(\mathbf{x}) = \frac{1}{(\det\Sigma)^{\frac{1}{2}}(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)}.$$

(i) Using definition II, and assuming Σ is positive definite, show that $E[\mathbf{X}] = \mu$ and $E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T] = \Sigma$. [Hint: By spectral theorem, a positive definite matrix can be split into a product orthogonal and diagonal matrices as follows: $\Sigma = U\Lambda U^T$, where U is an orthogonal matrix, that is, $UU^T = I_n$ and Λ is a diagonal matrix with entries along the diagonal equal to the eigenvalues of Σ. Using this, show that $\mathbf{Y} = U^T(\mathbf{X} - \mu) \sim \mathcal{N}(0, \Lambda)$. Then use $E[\mathbf{X}] = E[U\mathbf{Y} + \mu]$, and $E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T] = E[(U\mathbf{Y})(U\mathbf{Y})^T]$.]

(ii) Using definition I, show that any subset of $\mathbf{X}$ is also multivariate Gaussian.

(iii) If two variables X_i and X_j of a multivariate Gaussian vector are uncorrelated, that is, $E[(X_i - \mu_i)(X_j - \mu_j)] = 0$, then show that X_i and X_j are independent. This is a special property of multivariate Gaussian distributions and is not true of general multivariate distributions.