

# CSC 310 Information Theory

## Preparation for Test #1

Periklis Papakonstantinou

University of Toronto

The first test evaluates your knowledge on the following prerequisite mathematical material.

- Discrete Probability and Enumeration Principles
- Linear Algebra and elements of Abstract Algebra
- Elementary knowledge of Calculus

Revise the material carefully and (before anything else) resolve any questions regarding notation and terminology. You should know both the basic definitions, the intuition behind definitions and proofs, and be competent at a technical level (technical=proving skill and expression). Clearly, we require, and you will be tested on, much less and also simpler material than the prerequisite courses. However, we demand (i) *solid* foundation on the basic concepts, (ii) you have clarified common misconceptions, pitfalls and fallacies, (iii) you understand how to read and write clear mathematical proofs and in particular related to the above areas.

The given review/preparation time is small. Avoid getting lost in piles of texts. Start your study from the provided list of *simple* intuitive and technical questions. The intuitive questions are to the point, but since they are intuitive it is possible for different answers to be correct (do a best effort to answer them). Clearly, the technical questions form just a sparse sample - far from being complete/adequate.

## 1 Necessary elements of Calculus

- We assume familiarity with the concepts of monotonicity, differentiation, and extreme values.
- Define  $\lim_{n \rightarrow \infty} f(n) = \alpha$ ,  $\alpha \in \mathbb{R}$ .
- Show that  $(1 - \frac{1}{n})^n \leq \frac{1}{e}$ .
- State the Mean Value Theorem.
- For which values of  $x$  is it true that  $(1 - x) \leq e^{-x}$ ? Explain.

## 2 Discrete Probability and Enumeration Principles

Just focus on Discrete Probability. No need to study measure-theoretic probability.

**Intuitive & Definitional questions.** Do a best effort to give a satisfactory answer.

1. When we write  $\Pr(A)$ , to what the symbols  $\Pr(\cdot)$  and  $A$  correspond to? Does the notation  $\Pr(A)$  suffice to describe what is going on?
2. Let  $A, B$  be two events over a uniform probability space  $\Omega$ . Does the notation  $\Pr(A|B)$  correspond to a probability? Is  $A|B$  an event?
3. Define  $\Pr(A|B)$ . Intuitively which probability space we could have associated with  $\Pr(A|B)$ .
4. Is a random variable a variable? If not what is it?
5. What is the intuitive relation (if any) between a random variable and events of a probability space?
6. Why random variables are interesting objects? Why do we care to define them?
7. What is a moment of a random variable  $X$ ? Define the expectation  $\mathbb{E}(X)$  of a variable and its variance  $\text{Var}(X)$ .
8. What is the difference between the expectation and the variance? Which of the two has linearity properties? Why do you think that linearity is important?
9. Intuitive claim:  
*the term “expectation” in some cases may be misleading, since the expected value of an experiment may have nothing to do to what “we expect” from the experiment.*  
Do you agree or disagree with this claim. Justify through examples.
10. Give examples of random variables where (i)  $X$  is concentrated around  $\mathbb{E}(X)$  and (ii) examples where it's not concentrated  $\mathbb{E}(X)$ .
11. State the Markov's and Chebychev's inequalities. What's their difference?
12. Consider an experiment associated with a probability space  $\Omega$ . We say that two independent executions of the experiment formally correspond to the product space  $\Omega \times \Omega$ . Define precisely  $\Omega \times \Omega$  as a probability space. Intuitively, why do independent executions of the experiment are associated with the product space?
13. State Bayes Rule and intuitively explain what it says.
14. Are the terms *mutually exclusive events* and *independent events* related?
15. State the Law of Large Numbers and give an intuitive explanation.
16. Give examples clarifying further all of the above questions.

Here, we don't provide questions from enumeration principles. Enumeration principles will be tested. Please consult the relevant handout.

**Technical questions.** As mentioned, we only consider finite probability spaces.

1. Show that for any two events  $A, B \subseteq \Omega$  of the probability space  $\Omega$  we have that  $\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$ . Is this inequality strict? Does the inequality hold with equality in case of independent events? Does it hold with equality in case of disjoint events? Provide both a technical and an intuitive explanation.
2. Let  $X$  be a random variable and  $a, b \in \mathbb{R}$ . Show that  $\text{Var}(aX + b) = a^2\text{Var}(X)$ . What does this identity intuitively state?
3. Let  $A_1, A_2, \dots, A_n \subseteq \Omega$  be  $n$  events. Show the *union bound*  $\Pr(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \Pr(A_i)$ .
4. Suppose that the year has 365 days and consider the following experiment. Consider a class of students 30 each with a birthday chosen uniformly at random. What is the probability that 2 students were born at the same day? In your calculations present in detail the probability spaces.
5. Consider the experiment where we throw (uniformly at random)  $m$  balls to  $n$  bins. (i) Show that  $m = \sqrt{2n} + 1$  balls are needed such that the probability that one bin contains two balls becomes at least  $1 - \frac{1}{e}$ . (ii) What happens to this probability when  $m$  grows larger than  $\sqrt{2n} + 1$ ?
6. Prove Markov's inequality.
7. Prove Chebyshev's inequality.
8. Let  $X, Y$  be independent random variables. (i) Show that  $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ . (ii) Is this in general (i.e. not necessarily for independent r.v) true?
9. Let  $X, Y$  be independent random variables. (i) Show that  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ . (ii) Is this in general true?
10. Let  $X_1, \dots, X_m$  be independent random variables. Let  $X = \sum_{i=1}^m X_i$ . Show that  $\text{Var}(X)^2 = \sum_{i=1}^m \text{Var}(X_i)^2$ .
11. Let  $S, T$  be two disjoint subsets of  $U$  such that  $|S| = |T| = n$ . Suppose that we select a random  $R \subseteq U$  by independently sampling each element of  $U$  with probability  $p$ . We say that  $R$  is good if (i)  $R \cap S = \emptyset$  and (ii)  $R \cap T = \emptyset$ . Show that for  $p = 1/n$  the probability that  $R$  is good is larger than some positive constant.
12. Coupon collector's problem: suppose that we have  $n$  types of coupons and we perform the following experiment. At each independent trial we choose a coupon at random (which it can be of any of the  $n$  types). The goal is to end up having at least one coupon of each type. Show that the number of trials  $X$  (in order to have a coupon of each type) has expected value  $\mathbb{E}(X) = O(n \log n)$ .  
Remark: in coupon's collector we can actually show that  $X$  is concentrated around  $\mathbb{E}(X)$ .

### 3 Linear Algebra and elements of Abstract Algebra

#### Intuitive & Definitional questions

1. We expect that you understand elementary concepts such as linear independence and that you feel at home with performing matrix computations, solving linear systems, doing Gauss eliminations, computing determinants, and other elementary matrix computations. If you feel you should remind yourself these things then do it before anything else.
2. Here is an intuitive statement: “linearity an important concept all over mathematics, engineering and applied sciences”. Do you agree with this, and if yes to what extent? Give concrete examples.
3. Consider the following three algebraic objects: (i) group, (ii) field, (iii) vector space. How does one relate to the other? Can we say which one is richer (has more algebraic structure)? Give examples for all three objects in case they have infinitely and finitely many elements.
4.  $\mathbb{Z}_2 := \{0, 1\}$  equipped with ( mod 2) addition and multiplication is a field, and in particular it is an abelian group under addition. Can we say the same thing for  $\mathbb{Z}_3 = \{0, 1, 2\}$  or  $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ ?
5. Optional question: How is the finite field of 3 elements defined?
6. Is the set of integers  $\mathbb{Z}$  a group under addition? Is it a field?
7. Define a  $K$ -vector space<sup>1</sup>.
8. Define the sum and the direct sum of two vector spaces. Define the direct product of two vector spaces. What are the relations (if any) among these concepts.
9. Define what is a linear map (or linear function, or linear operator). What is the precise relation between a linear map and a matrix (define carefully the vector spaces of reference, fix bases etc - you can assume we are only interested in finite dimensional spaces).
10. How does the kernel of a matrix relate to the way the associated linear map works?
11. What is the precise relation between matrix multiplication and composition of linear maps. Give examples.
12. Define the determinant of a square matrix over  $\mathbb{Z}_2^{n \times n}$ . Which interpretations/ characterizations/definitions the determinant do you know?
13. What is an isomorphism between vector spaces?
14. What is a projection operator. Give an example. Is a projection operator always singular?
15. What can be inferred from the row-reduced Echelon form of a matrix?
16. Name some important properties of stochastic matrices. Give examples, and prove the properties.
17. Define the Hadamard matrix.
18. Define the scalar (dot) product on a vector space  $V$ . When do we say that  $u, v \in V$  are orthogonal to each other (or perpendicular)?

---

<sup>1</sup>Here  $K$  is a field; e.g.  $K = \mathbb{R}$ . A  $K$ -vector space or a vector space over  $K$ , is a vector space where the field of scalars is  $K$ . For example,  $\mathbb{R}^3$  is a  $\mathbb{R}$ -vector space.

19. State the Cauchy - Schwartz inequality. Intuitively, explain the importance of the inequality.

*Remark:* although eigenvalues and eigenvectors are basic to linear algebra we won't use them in CSC310. As optional intuitive questions you may consider: What is a symmetric matrix? State the spectral theorem e.g. for real symmetric matrices and explain its importance.

**Technical questions.** All vector spaces are finite dimensional.

1. Let  $V$  be a vector space, and let  $u_1, \dots, u_n \in V$ . Define  $W$  to be the set of all linear combinations of  $u_1, \dots, u_n$ . Show that  $W$  is a subspace of  $V$ .
2. Show that the set  $V$  whose elements are all continuous functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a vector space. (intuitive remark: make sure you understand what it means of an element of  $V$  to be a function - as far as it concerns the definition of vector space we don't really care how "complicated" each element of the vector space is.)
3. Let  $U, W$  be subspaces of  $V$ . (i) Show that  $U + W := \{u + w | u \in U, w \in W\}$  is a subspace of  $V$ . (ii) How about  $U \cap W$  or (iii)  $U \cup W$ ?
4. Show that the vectors  $(1, 1), (-3, 2) \in \mathbb{R}^2$  are linearly independent.
5. Show that the  $2 \times 2$  matrices (i.e.  $K^{2 \times 2}$ ) form a  $K$ -vector space. What is the dimension of this space? Give a basis.
6. Let  $F$  be a field and  $GL_n(F) = \{A | A \in F^{n \times n}, \det(A) \neq 0\}$ , i.e.  $GL_n(F)$  is the set of non-singular  $n \times n$  matrices with elements from  $F$ . If  $|F| = q < \infty$  then show  $|GL_n(F)| = (q^n - 1)(q^n - q) \dots (q^n - q^{n-1})$ .
7. For square matrices  $A, B$  show<sup>2</sup> that (i)  $\det(A) = \det(A^T)$  and (ii)  $\det(AB) = \det(A)\det(B)$ .
8. Let  $H$  be an  $n \times n$  Hadamard matrix. Show that  $(\det(H))^2 = n^n$ . (Hint: determine  $HH^T$  first.)
9. (i) Let  $V$  be a  $K$ -vector space, and let  $\mathcal{B}_V = \{u_1, \dots, u_n\}$  be a basis of  $V$ . Show that  $w_1, \dots, w_{n+1} \in V$  is a set of linearly dependent vectors. (ii) Conclude that the dimension  $\dim V$  of a vector space is well-defined (i.e. if  $\mathcal{B}_1, \mathcal{B}_2$  are two bases of  $V$  then  $|\mathcal{B}_1| = |\mathcal{B}_2|$ ).
10. Let  $E \in K^{n \times n}$  be a matrix of full rank  $\text{rank}(E) = n$ , and  $A \in K^{m \times n}$  a matrix. Show that  $\text{rank}(AE) = \text{rank}(A)$ .
11. Define elementary matrices and Gauss Elimination. Show that Gauss Elimination preserves the rank of the matrix.
12. Let  $U$  be a subspace of  $V$ . Then, there exists  $W$  subspace of  $V$  such that  $U \oplus W = V$ , where  $\oplus$  denotes direct sum. (intuitively explain what does this theorem say - give examples).
13. (i) Let  $V = U \oplus W$ . Then,  $\dim V = \dim U + \dim W$ . (ii) For any two vector spaces  $U, W$  show that  $\dim(U \times W) = \dim U + \dim W$ .

---

<sup>2</sup>For a matrix  $A$  we denote by  $A^T$  its transpose. Note that it's also common to denote the transpose by  ${}^t A$ .

14. Let  $A$  be strictly upper triangular matrix, i.e.

$$A = \begin{pmatrix} 0 & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & 0 & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Show that  $A^n = 0$ . Intuitively: what does this tell us for the composition of the corresponding linear operator? Give an example.

15. Let  $V$  be a vector space and the function  $T_u : V \rightarrow V$ , such that  $T_u(v) = u + v$ . (i) Show that  $T_u$  is a bijection. (ii) Show that  $T_u$  is *not* in general a linear mapping. (intuitive remark: we think of linear maps/transformations as being very simple functions which “stretch/rotate” vectors, but in general can’t do more things. Here, too  $T_u$  is in some sense simple, but it violates linearity. Which property of linearity such a simple function violates?)
16. Let  $V$  and  $W$  be  $K$ -vector spaces. Let  $\mathcal{B}_V = \{u_1, \dots, u_n\}$  be a basis of  $V$  and  $\mathcal{B}_W = \{w_1, \dots, w_n\}$  be a basis of  $W$ . Then, there exists unique linear mapping  $T : V \rightarrow W$  such that: (i)  $T(u_1) = w_1, \dots, T(u_n) = w_n$  and (ii)  $T(a_1u_1 + \dots + a_nu_n) = a_1w_1 + \dots + a_nw_n$ , for every  $a_i \in K$ .
17. Let  $V, W$  be vector spaces. Let  $T : V \rightarrow W$  be a linear mapping. Then,  $\dim V = \dim \ker T + \dim \text{Im} T$ . Intuitively: in this statement at a first glance the space  $W$  seems to have nothing to do with the statement - what is the role of  $W$ ?
18. Let  $T : K^n \rightarrow K^m$  be a linear map. Then, there exists unique vector  $A \in K^n$  such that  $T = T_A$ , i.e.  $L(X) = A^T X$ . More generally, consider the linear map  $T : K^n \rightarrow K^m$ , and determine the matrix  $A$  associated with  $T$ .
19. **Notation:** let  $V, W$  be  $K$ -vector spaces with bases  $\mathcal{B}_V = \{v_1, \dots, v_n\}$  and  $\mathcal{B}_W = \{w_1, \dots, w_m\}$ . Then, every  $v \in V$  is uniquely written as  $v = x_1v_1 + \dots + x_nv_n$ ,  $x_i \in K$ . That is,  $V$  is isomorphic to  $K^n$  with isomorphism  $(x_1, \dots, x_n) \mapsto x_1v_1 + \dots + x_nv_n$ ; and similarly for  $W$ . Let  $f : V \rightarrow W$  be a linear map. Using the above isomorphism we can interpret  $f$  as a  $K^n \rightarrow K^m$  map, with associated matrix  $\mathcal{M}_{\mathcal{B}_W}^{\mathcal{B}_V}(f)$ . The notation  $\mathcal{M}_{\text{to vector space}}^{\text{from vector space}}$  (linear map). Then, given a vector  $v \in V$  we can express it uniquely as a column vector  $X_{\mathcal{B}_V}(v) \in K^n$  such that  $X_{\mathcal{B}_W}(f(v))$  (i.e.  $f(v)$  identified as a vector in  $K^m$ ) can be calculated as follows:  $\mathcal{M}_{\mathcal{B}_W}^{\mathcal{B}_V}(f)X_{\mathcal{B}_V}(v) = X_{\mathcal{B}_W}(f(v))$ .
- To show:** Let  $V, W, U$  be vector spaces, and let  $\mathcal{B}, \mathcal{B}', \mathcal{B}''$  be corresponding bases. Let  $f : V \rightarrow W$  and  $g : W \rightarrow U$  be linear maps. Show that,  $\mathcal{M}_{\mathcal{B}''}^{\mathcal{B}'}(g)\mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(f) = \mathcal{M}_{\mathcal{B}''}^{\mathcal{B}}(g \circ f)$ .
20. Let  $V$  be a vector space,  $\mathcal{B}$  a basis of  $V$  and  $\text{id}$  be the identity map. (i) Show that  $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\text{id}) = I$ . (ii) For two distinct bases  $\mathcal{B}, \mathcal{B}'$  of  $V$ , is it true that  $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(\text{id}) = I$ ?
21. Let  $f : V \rightarrow W$  be an invertible linear map. Show that for every choice of bases  $\mathcal{B}_V, \mathcal{B}_W$ :  $\mathcal{M}_{\mathcal{B}_V}^{\mathcal{B}_W}(f^{-1})\mathcal{M}_{\mathcal{B}_W}^{\mathcal{B}_V}(f) = I$ .
22. Let  $V$  be a finite-dimensional  $\mathbb{R}$ -vector space with a positive definite scalar product. Let  $W$  be a subspace of  $V$  and let  $W^\perp$  be the set of all vectors perpendicular to  $W$ . Show (i) that

$W^\perp$  is a vector space (and in particular a subspace of  $W$ ), and (ii)  $\dim W + \dim W^\perp = \dim V$ . Intuitively explain how  $W$  and  $W^\perp$  look like. Give examples.