Mathematical Foundations of Data Science
(COMS4055A, COMS7058A)
Class Test 1

13 March 2024, 14h15–16h15, JD du Plessis

Name: ______________________ Row: ______ Seat: ______ Signature: __________

Student Number: ______________ ID Number: ______________

Instructions

- Answer all questions in pen. **Do not write in pencil.**
- This test consists of 3 pages. Ensure that you are not missing any pages.
- You may bring in a single A4 page of notes.
- You are allocated 2 hour to complete this test.
- Ensure your cellphone is switched off.
- You may use a calculator during the test.
- Round off to an appropriate number of decimal places and simplify your answers fully.
- Please manage your time appropriately. Some questions are difficult but worth few marks, do not linger on a difficult question if you have not answered others.

Question 1 Linear Algebra [60 Marks]

1. Using the digits 1, 2, 3, 4, 5 and 6:

   (a) How many three digit numbers can be made if digits may be repeated? [2]
   $6 \cdot 6 \cdot 6 = 216$

   (b) How many three digit numbers can be made if digits may **not** be repeated? [2]
   $6 \cdot 5 \cdot 4 = 120$

   (c) How many three digit numbers can be made if digits may **not** be repeated and must go from smallest to largest? [2]
   $\binom{6}{3} = 20$, equivalently you could observe that exactly one sixth of the above 120 numbers have digits in increasing order.

2. Many books and articles analyze the strategies of highly successful entrepreneurs. These sources often highlight qualities such as risk-taking, persistence, and unconventional thinking as key factors in achieving success.

   What potential flaw exists in drawing conclusions about success based only on studying well-known, successful entrepreneurs? [4]
   Survivor bias. Perhaps those traits are also shared by entrepreneurs who failed.

3. A school has five classes with the following class sizes: $\{10, 20, 30, 40, 50\}$ students. Students are randomly surveyed and asked how many students are in their class. The school notices that the average reported class size is larger than the true average class size when computed from school records.

   (a) Explain why the students' reported average class size is larger than the true average class size. [2]
   Size bias. Students from more populous classes will be sampled more often by virtue of there being more of them.

   (b) Compute the size of the average class based on the five class sizes given. [2]
   $\frac{10+20+30+40+50}{5} = 30$

   (c) If a student is picked at random and reports their class size, what is the expected class size they experience? [2]
   $\frac{10^2+20^2+30^2+40^2+50^2}{10+20+30+40+50} = 36.666666\ldots$

4. Let X be a continuous random variable with probability density function (pdf) given by:

$$f(x) = \begin{cases} c(4-x)^2, & 0 \leq x \leq 3, \\ 0, & \text{otherwise.} \end{cases}$$

where c is a constant.

(a) Find the constant c that makes $f(x)$ a valid probability density function. [3]

$$\begin{aligned}
1 &= \int_0^3 c(4-x)^2 dx \\
\frac{1}{c} &= \int_0^3 16 - 8x + x^2 dx \\
\frac{1}{c} &= 16x - 4x^2 + \frac{x^3}{3}|_0^3 \\
\frac{1}{c} &= 21 \\
c &= \frac{1}{21}
\end{aligned}$$

(b) Compute the expected value $E[X]$ of the distribution. [2]

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^3 \frac{x}{21}(4-x)^2 dx \\
&= \frac{1}{21}\int_0^3 16x - 8x^2 + x^3 dx \\
&= \frac{1}{21}(8x^2 - \frac{8}{3}x^3 + \frac{x^4}{4}|_0^3) \\
&= \frac{81}{84}
\end{aligned}$$

(c) Compute the variance Var(X) of the distribution. [3]

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^3 \frac{x^2}{21}(4-x)^2 dx \\
&= \frac{1}{21}\int_0^3 16x^2 - 8x^3 + x^4 dx \\
&= \frac{1}{21}(\frac{16}{3}x^3 - 2x^4 + \frac{x^5}{5}|_0^3) \\
&= \frac{153}{105}
\end{aligned}$$

$$\begin{aligned}
V(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{153}{105} - (\frac{81}{84})^2 \\
&= 0.52729591836
\end{aligned}$$

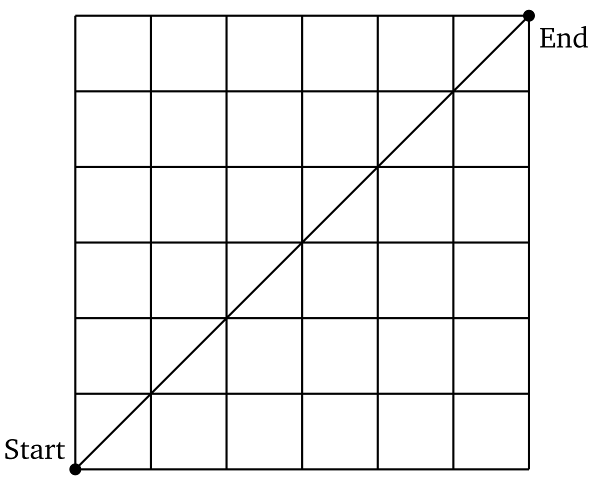
(d) Compute the third moment $E[X^3]$ of the distribution. [2]

$$\begin{aligned}\mathbb{E}[X^3] &= \int_0^3 \frac{x^3}{21}(4-x)^2dx \\ &= \frac{1}{21}\int_0^3 16x^3 - 8x^4 + x^5 dx \\ &= \frac{1}{21}(4x^4 - \frac{8}{5}x^5 + \frac{x^6}{6}|_0^3) \\ &= 2.7\end{aligned}$$

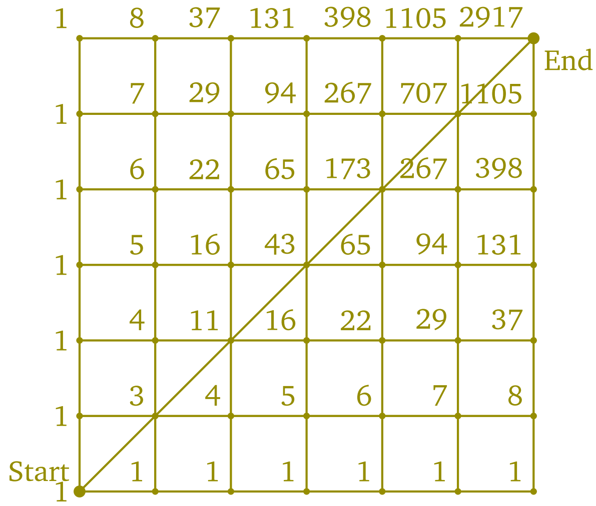
5. A fair coins is tossed 12 times and the number of heads is recorded as X. Compute:
   (a) Compute $\mathbb{E}[X]$ [2]
   $\mathbb{E}[X] = np = 12 \cdot \frac{1}{2} = 6$
   (b) Compute $V(X)$ [2]
   $V(X) = npq = 12 \cdot \frac{1}{2} \cdot \frac{1}{2} = 3$
   (c) $\mathbb{P}(X >= 10)$ [3]

$$\begin{aligned}\mathbb{P}(X >= 10) &= \mathbb{P}(X = 10) + \mathbb{P}(X = 11) + \mathbb{P}(X = 12) \\ &= \frac{\binom{12}{10}}{2^{12}} + \frac{\binom{12}{11}}{2^{12}} + \frac{\binom{12}{12}}{2^{12}} \\ &= \frac{66}{2^{12}} + \frac{12}{2^{12}} + \frac{1}{2^{12}} \\ &= 0.01928710937\end{aligned}$$

6. How many ways can one go from the start to the end in the picture below if you are only move along the drawn in lines and may only move towards to end (up, right or diagonally up and right. Leftwards and downwards movement are prohibited)? [4]



We compute recursively.



7. Recall that a random variable X is said to follow a **Gamma distribution** with parameters $\alpha > 0$ and $\lambda > 0$ if it's probability density function (pdf) is given by

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0.$$

Suppose that $X_1 \sim \text{Gamma}(\alpha_1, \lambda)$ and $X_2 \sim \text{Gamma}(\alpha_2, \lambda)$ are independent with α_1 and α_2 integer. Define the sum $S = X_1 + X_2$.

Find the distribution of S [3]

$Gamma(\alpha_1 + \alpha_2, \lambda)$ X_1 and X_2 are the sums of α_1 and α_2 exponential with parameter λ so their sum is the sum of $\alpha_1 + \alpha_2$ such variables.

8. Let X be distributed exponetial with mean $\lambda = 2$. Compute: $\mathbb{P}(3 < X < 6)$ [3]

Typo here. Typicaly $\frac{1}{\lambda}$ is the mean. Marks were awarded weather you used $\lambda = 2$ pr $\lambda = \frac{1}{2}$.

$$\begin{aligned} \mathbb{P}(3 \leq X \leq 6)) &= \mathbb{P}(X \geq 3) - \mathbb{P}(X \geq 6) \\ &= e^{-3\lambda} - e^{-6\lambda} \end{aligned}$$

9. Three fair six sided dice are rolled giving values X_1 , X_2 and X_3. Let $Y = max(X_1, X_2, X_3)$ be the largest of these rolls. Compute:

   (a) $\mathbb{P}(Y = 6)$ [2]

   This is equalivent to asking for the probability of at least one 6, which is $1 - (\frac{5}{6})^3 = \frac{91}{216}$. Here the $(\frac{5}{6})^3$ term represents the probability of getting 3 rolls which are 5 or below.

(b) $\mathbb{P}(Y = 5)$ [2]
By similar reasoning $(\frac{5}{6})^3 - (\frac{4}{6})^3 = \frac{61}{216}$

(c) $\mathbb{E}[Y]$ [4]
Use similar reasoning to compute the probabilities of $Y = 1, 2, 3, 4$ and compute:

$$\mathbb{E}[Y] = 6 \cdot \frac{91}{216} + 5 \cdot \frac{61}{216} + 4 \cdot \frac{37}{216} + 3 \cdot \frac{19}{216} + 2 \cdot \frac{7}{216} + 1 \cdot \frac{1}{216} = 4.95833333333$$

10. Six children are at a birthday party. They each write down their name and put it in a hat. The names in the hat are then well shuffled and each child draws out one of the pieces of paper, emptying the hat. Compute: Let X be the random varible counting the number of children who received their own name. Compute:

(a) $\mathbb{E}[X]$. The expected number of children who receive their own name. [2]
1, each child has a one in six chance of drawing their own name. We then apply linearity of expectation.

(b) $\mathbb{P}(X = 0)$. The probability that no child receives their own name [3]
Derangements! $D_6 = 6!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!}) = 265$

(c) $\mathbb{P}(X = 2)$. The probability that exactly two children receive their own name. [4]
$\binom{6}{2} D_4 = 15 \cdot 9 = 135$ We first choose two of the 6 to receive their own names and the remainign 4 must be a derrangement.