DATA SCIENCE **Stochastic Methods**	Name:________________________
January 21, 2021 Prof. Marco Ferrante	Student number:____________________

Problem 1. [10] A box of chocolates contains 10 milk chocolates and 5 dark chocolates. Every time that we eat a chocolate, we take this form the box randomly and independently of the previous choices. Let X denotes the (random) number of dark chocolates still in the box after we have eaten 3 chocolates.

(i) Define the support and evaluate the discrete density of the random variable X; 4

(ii) Compute $\mathbb{P}[X > 3|X \leq 4]$; 3

(iii) Compute the moment generating function of X. 3

(i) $X \in \{2, 3, 4, 5\}$

$\mathbb{P}[X=2] = \frac{5}{15} \cdot \frac{4}{14} \cdot \frac{3}{13}$, $\mathbb{P}[X=3] = 3 \cdot \frac{10}{15} \cdot \frac{5}{14} \cdot \frac{4}{13}$, $\mathbb{P}[X=4] = 3 \cdot \frac{10}{15} \cdot \frac{9}{14} \cdot \frac{5}{13}$

$\mathbb{P}[X=5] = \frac{10}{15} \cdot \frac{9}{14} \cdot \frac{8}{13}$

(ii) $\mathbb{P}[X>3 | X \leq 4] = \frac{\mathbb{P}[3 < X \leq 4]}{\mathbb{P}[X \leq 4]} = \frac{\mathbb{P}[X=4]}{1 - \mathbb{P}[X=5]}$

(iii) $m_X(t) = \mathbb{E}[e^{tX}] = e^{2t}\,\mathbb{P}[X=2] + e^{3t}\,\mathbb{P}[X=3]$

$+ e^{4t}\,\mathbb{P}[X=4] + e^{5t} \cdot \mathbb{P}[X=5]$

Problem 2. [12] Let $X_1, \ldots, X_n$ be independent, positive random variables, with common density

$$f(x) = \frac{1}{2} x^2 e^{-x} \, \mathbf{1}_{(0,+\infty)}(x)$$

(i) Compute the distribution of X_1, i.e. $\mathbb{P}[X_1 \le x]$ for any $x \in \mathbb{R}$; 3

(ii) Compute $m(t) = \mathbb{E}[e^{tX}]$ and $\mathbb{E}[X] = 3$. 3

(iii) Find the Chernoff lower tail estimate 6

$$\mathbb{P}[\widehat{X}_n \le 3 - \varepsilon] \le e^{-n\frac{\varepsilon^2}{6}}$$

Hint: use the inequality $\log(1-x) \le -x - x^2/2$.

(i) $X_1 \sim \Gamma(3,1)$ $\qquad \mathbb{P}[X_1 \le x] = \int_0^x \frac{1}{2} y^2 e^{-y} dy$

$$= -\frac{1}{2} y^2 e^{-y} \Big|_0^x + \int_0^x y e^{-y} dy$$

$$= -\frac{1}{2} x^2 e^{-x} - y e^{-y} \Big|_0^x + \int_0^x e^{-y} dy$$

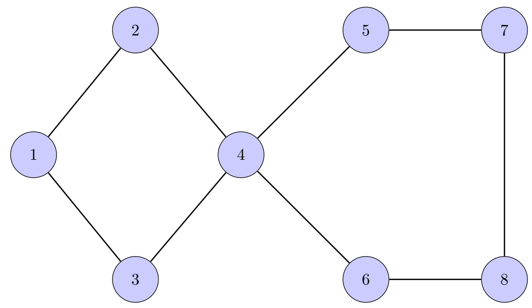
$$= -\frac{1}{2} x^2 e^{-x} - x e^{-x} - e^{-x} + 1$$

(ii) $m(t) = \int_0^{+\infty} e^{ty} \frac{1}{2} y^2 e^{-y} dy = \frac{1}{(1-t)^3}$ $\quad t < 1, \quad +\infty \quad t \ge 1$

$m'(t) = 3 \frac{1}{(1-t)^4}$ $\qquad \mathbb{E}[X] = m'(t)\big|_{t=0} = 3$

(iii) $\mathbb{P}[\widehat{X}_n \le 3 - \varepsilon] \le e^{-n\, h(t)}$, where $h(t) =$

Problem 3. [12] Define a simple Random Walk $\{X_n, n \geq 0\}$ on the undirected graph:



(i) Find the invariant distribution. 4

(ii) Is this distribution reversible? 2

(iii) Starting from the state 1, how many steps are needed on average to go back to the state 1? 3

(iv) Starting from state 1, what is the probability that we arrive for the first time to state 5 before visiting for the first time the state 6? 3

(i) Let σ_i = # neighbour nodes of i^{th} node

$$\sigma_1 = \sigma_2 = \sigma_3 = \sigma_5 = \sigma_6 = \sigma_7 = \sigma_8 = 2, \quad \sigma_4 = 4$$

$\sum_{i=1}^{8} \sigma_i = 18$. The invariant distrib. is π, with $\pi_i = \frac{2}{18}$ $\forall i \neq 4$

$\pi_4 = 4/18$

(ii) YES

(iii) $m_1 = \frac{1}{\pi_1} = \frac{18}{2} = 9$

(iv) We can define a bijection from the paths that start in 1 and reach 5 before 6 with those that start in 1 and reach 6 before 5. Since with probability 1 we reach 5 or 6, then the probability to arrive to 5 before visiting 6 is equal to $1/2$.