Appello di Teoria dell'Informazione e Inferenza del 12 Luglio 2024

1.1 Quanti sono gli anagrammi di PERDERE comprese le parole senza senso?

1.2 Ricava le probabilità marginali e calcola $\mathbb{E}[X^2]$ e $\mathbb{E}[Y]$, se

$$
\begin{aligned}
p(X=1,Y=2) = p(X=2,Y=3) = p(X=3,Y=4) &= \frac{1}{27}\\
p(X=1,Y=3) = p(X=2,Y=2) = p(X=3,Y=2) &= \frac{1}{9}\\
p(X=1,Y=4) = p(X=2,Y=4) = p(X=3,Y=3) &= \frac{5}{27}
\end{aligned}
$$

da rivedere

1.3 Determina il valore di a per il quale la funzione

$$f(x) = \begin{cases} a & \text{se } x \in [0,1/2] \cup [1,3/2] \\ 0 & \text{altrimenti} \end{cases}$$

è una *pdf*. Calcola la *cdf* della variabile casuale continua X sottostante e producine il grafico.

1.4 Se $X_1, \ldots, X_n$ sono valori ottenuti campionando una variabile casuale X, spiega perché la media empirica $\sum_i X_i/n$ è una variabile casuale mentre il valore atteso $\mathbb{E}[X]$ non lo è.

2.1 Se $H(Y) = 2$ e $H(X,Y) = 5$ che cosa puoi dire di $H(X|Y)$ e di $H(Y|X)$ se $H(X) = 4$?

2.2 Date le equazioni di parità

bho

$$
\begin{aligned}
y_1[n] &= x[n] + x[n-1] + x[n-2]\\
y_2[n] &= x[n] + x[n-1]\\
y_3[n] &= x[n] + x[n-2]
\end{aligned}
$$

se $x[3] = 0$ e $x[4] = 0$ trova la codifica per il bit $x[5] = 1$ e calcolane la distanza di Hamming dalla tripletta 111.

2.3 Dati i 4 simboli a, b, c e d, per quale motivo una codifica C tale che $L_C(a) = 1$ e $L_C(b) = 2$, $L_C(c) = 2$ e $L_C(d) = 3$ non può essere di Huffman?

a studio **2.4** Trova gli intervalli associati dalla codifica aritmetica alle triplette 001 e 111 se $p(0) = 1/4$ e $p(1) = 3/4$.

3.1 Un cassetto contiene 4 monete che restituiscono *croce* con probabilità 1/8 e 2 monete che restituiscono *croce* con probabilità 1/4. Qual è la probabilità di ottenere *croce* lanciando una moneta pescata a caso? E quale quella di ottenere testa lanciandola nuovamente?

3.2 Hai due monete identiche per forma e peso. La prima restituisce *testa* con probabilità 1/2, la seconda con probabilità 1/10. Ne scegli una a caso e la lanci 4 volte ottenendo la sequenza $TCCC$. Calcola la verosimiglianza delle due possibili scelte.

3.3 Per quale motivo

$$\mathbf{P} = \begin{pmatrix} 0.5 & 0.5 \\ 0 & 1 \end{pmatrix}$$

non può essere la matrice di transizione di un processo di Markov regolare? Qual è la sua distribuzione stazionaria?

3.4 Se una matrice di transizione di un processo di Markov ha due distribuzioni stazionarie, perché non può avere la distribuzione limite?

1.1

$$\frac{7!}{3!\,2!} = \frac{5040}{6 \cdot 2} = 420$$

1.2

PARTIAMO A CALCOLARE LE MARGINALI

$$P(X=1): \frac{1}{27} + \frac{1}{9} + \frac{5}{27} = \frac{1+3+5}{27} = \frac{9}{27} = \frac{1}{3}$$

$$P(X=2): \frac{1}{27} + \frac{1}{9} + \frac{5}{27} = \frac{1}{3}$$

$P(X=3): \frac{1}{3}$

$P(Y=2): \frac{1}{27} + \frac{1}{9} + \frac{1}{9} = \frac{1+3+3}{27} = \frac{7}{27}$

$P(Y=3): \frac{1}{27} + \frac{1}{9} + \frac{5}{27} = \frac{1}{3}$

$P(Y=4): \frac{5}{27} + \frac{5}{27} + \frac{1}{27} = \frac{11}{27}$

$E[X^2]$

$$\left(1^2 \cdot \frac{1}{27}\right) + \left(2^2 \cdot \frac{1}{27}\right) + \left(3^2 \cdot \frac{1}{27}\right) + \left(1^2 \cdot \frac{1}{9}\right) + \left(2^2 \cdot \frac{1}{9}\right) + \left(3^2 \cdot \frac{1}{9}\right) + \left(1^2 \cdot \frac{5}{27}\right) + \left(2^2 \cdot \frac{5}{27}\right) + \left(3^2 \cdot \frac{5}{27}\right) =$$

$$= \frac{1}{27} + \frac{4}{27} + \frac{1}{3} + \frac{1}{9} + \frac{4}{9} + 1 + \frac{5}{27} + \frac{20}{27} + \frac{5}{3} = \frac{1+4+9+3+12+27+5+20+45}{27} = \frac{126}{27} \approx 4{,}66$$

$E[Y]$

UN MODO PIU' VELOCE E' QUELLO DI SFRUTTARE LE MARGINALI CALCOLATE PRIMA. QUINDI:

$$2 \cdot \frac{7}{27} + 3 \cdot \frac{1}{3} + 4 \cdot \frac{11}{27} = \frac{14}{27} + 1 + \frac{44}{27} = \frac{14+27+44}{27} = \frac{85}{27} \approx 3{,}14$$

1.3

TROVARE a

$$\int_0^{1/2} a + \int_1^{3/2} a = 1 \Rightarrow a\int_0^{1/2} 1 + a\int_1^{3/2} 1 = 1 \Rightarrow a\int_0^{1/2} x + a\int_1^{3/2} x = 1 \Rightarrow \frac{a}{2} + a\left[\frac{3}{2} - 1\right] = 1 \Rightarrow$$

$$\frac{a}{2} + \frac{a}{2} = 1 \Rightarrow a = 1$$

CALCOLO CDF

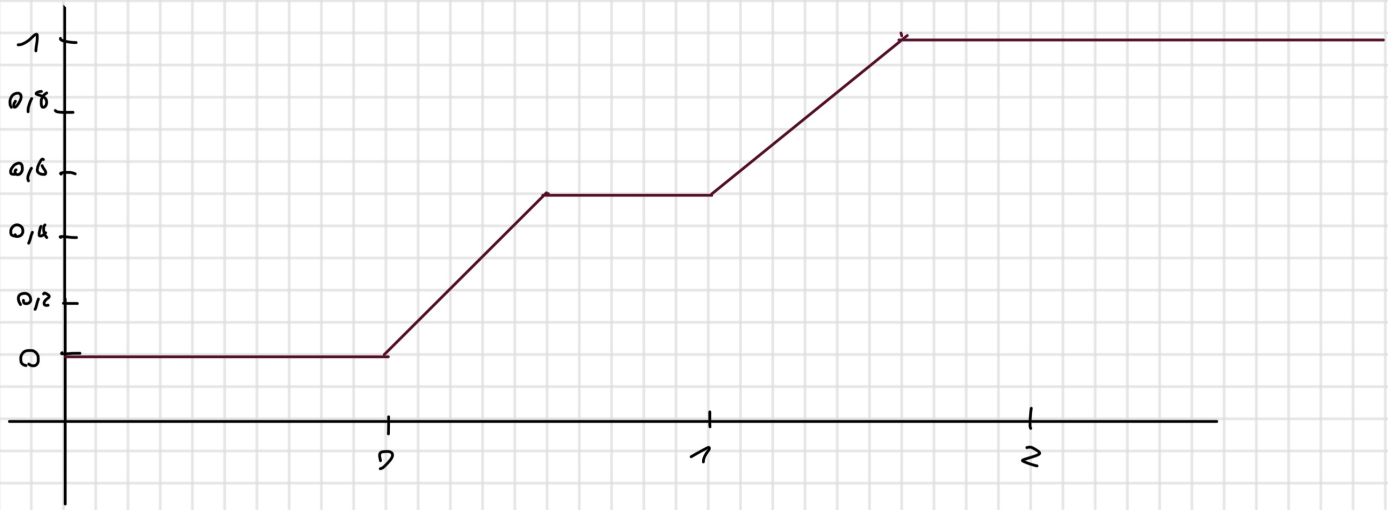
PER $x < 0$: $F(x) = 0$

PER $0 \leq x \leq 1/2$: $\int_0^x 1\, dt \Rightarrow x$

PER $1/2 < x < 1$: IN QUESTA ZONA LA DENSITA' E' 0. LA PROBABILITA' CUMULATA RIMANE COSTANTE AL VALORE RAGGIUNTO PRECEDENTEMENTE QUINDI $F(x) = 1/2$

PER $1 \leq x \leq 3/2$: $\int_1^x x + \frac{1}{2} \Rightarrow x - 1 + \frac{1}{2} \Rightarrow x - \frac{1}{2}$

PER $x > \frac{3}{2}$: 1



1.4

LA MEDIA EMPIRICA E' UNA FUNZIONE DI VARIABILI CASUALI, PERCIO' E' UNA VARIABILE CASUALE.

IL VALORE ATTESO NON DIPENDE DAL RISULTATO DEL SINGOLO EVENTO MA DALLA DISTRIBUZIONE DI X.

2.2

$y_1[5] = x[5] + x[4] + x[3] = 1+0+0 = 1$

$y_2[5] = x[5] + x[4] = 1$

$y_3[5] = x[5] + x[3] = 1$

DISTANZA HAMMING: 0 IN QUANTO 111 = 111

2.3

LA CODIFICA NON PUO' ESSERE DI HUFFMAN PERCHE' PRESENTA UN SOLO SIMBOLO CON LUNGHEZZA MASSIMA. IN UN ALBERO DI HUFFMAN BINARIO LE FOGLIE DI PROFONDITA' MASSIMA DEVONO COMPARIRE IN NUMERO PARI, CONDIZIONE CHE QUI NON E' SODDISFATTA.

2.4

STRINGA 001

$-x_1 = 0$

$a_1 = a_0 + L_0 \cdot CDF(-1) = 0$

$b_1 = a_0 + L_0 \cdot CDF(0) = 0 + 1 \cdot \frac{1}{4} = \frac{1}{4}$

$\Phi[0, \frac{1}{4})$ $\quad L_1 = \frac{1}{4}$

$-x_2 = 0$

$a_2 = a_1 + L_1 \cdot CDF(-1) = 0$

$b_2 = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$

$\Phi[0, \frac{1}{16})$ $\quad L_2 = \frac{1}{16}$

$x_3 = 1$

$$a_3 = a_2 + L_2 \cdot CDF(0) = \frac{1}{16} \cdot \frac{1}{4} = \frac{1}{64}$$

$$\beta_3 = \frac{1}{16}$$

$$\left[\frac{1}{64}, \frac{1}{16}\right)$$

-> Trace 111

$-x_1 = 1$

$$a_1 = a_0 + L_0 \cdot CDF(0) = \frac{1}{4}$$

$$\beta_1 = 1$$

$$\left[\frac{1}{4}, 1\right) \qquad L_1 = 1 - \frac{1}{4} = \frac{3}{4}$$

$-x_2 = 1$

$$a_2 = \frac{1}{4} + \left(\frac{3}{4} \cdot \frac{1}{4}\right) = \frac{1}{4} + \frac{3}{16} = \frac{4+3}{16} = \frac{7}{16}$$

$$\beta_2 = \frac{1}{4} + \frac{3}{4} = 1$$

$$\left[\frac{7}{16}, 1\right) \qquad L_2 = 1 - \frac{7}{16} = \frac{9}{16}$$

$-x_3 = 1$

$$a_3 = a_2 + L_2 \cdot CDF(0) = \frac{7}{16} + \left(\frac{9}{16} \cdot \frac{1}{4}\right) = \frac{7}{16} + \frac{9}{64} = \frac{28+9}{64} = \frac{37}{64}$$

$$\beta_3 = a_2 + L_2 \cdot CDF(1) = \frac{7}{16} + \frac{9}{16} = \frac{16}{16} = 1$$

3.1

$$P(C_1 | A) = \frac{1}{8} \quad P(A) = \frac{4}{6} \qquad P(C_1 | B) = \frac{1}{4} \quad P(B) = \frac{2}{6}$$

$$P(C_1) = \frac{1}{8} \cdot \frac{4}{6} + \frac{1}{4} \cdot \frac{2}{6} = \frac{1}{12} + \frac{1}{12} = \frac{2}{12} = \frac{1}{6}$$

$$P(A | C_1) = \frac{P(C_1 | A) \cdot P(A)}{P(C_1)} = \frac{1/12}{1/6} = \frac{1}{2}$$

$$P(B | C_1) = 1 - P(A | C_1) = \frac{1}{2}$$

$$P(T_2|C_1) = P(T_2|A) \cdot P(A|C_1) + P(T_2|B) \cdot P(B|C_1)$$

$$P(T_2|A) = 1 - P(C_1|A) = 1 - 1/8 = 7/8$$

$$P(T_2|B) = 1 - P(C_1|B) = 1 - 1/4 = 3/4$$

$$P(T_2|C_1) = \frac{7}{8} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} = \frac{13}{16}$$

3.2

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

CASO $p = 1/2$

$$\binom{4}{1} \cdot 1/2^1 \cdot (1 - 1/2)^3 \Rightarrow 4 \cdot \frac{1}{2} \cdot \frac{1}{8} = 2 \cdot \frac{1}{8} = \frac{1}{4} = 0.25$$

CASO $p = 1/10$

$$\binom{4}{1} \cdot \frac{1}{10} \cdot \left(1 - \frac{1}{10}\right)^3 \Rightarrow 4 \cdot \frac{1}{10} \cdot \left(\frac{10-9}{10}\right)^3 \Rightarrow \frac{4}{10} \cdot \frac{729}{1000} = \frac{2916}{10'000} \approx 0.29$$

3.3

NON E' REGOLARE IN QUANTO APPARE UNO 0. PER ESSERE REGOLARE DOVREMMO AVERE $P_{ij} > 0$

$$\begin{cases} \pi_1 = 0.5\pi_1 + 0\pi_2 \\ \pi_2 = 0.5\pi_1 + 1\pi_2 \end{cases}$$ CON VINCOLO $\pi_1 + \pi_2 = 1$

$$0.5\pi_1 = 0 \Rightarrow \pi_1 = 0$$

SOSTITUIAMO NEL VINCOLO $\rightarrow 0 + \pi_2 = 1 \Rightarrow \pi_2 = 1$

LA DISTRIBUZIONE STAZIONARIA E': $\pi[0, 1]$

3.4

LA PRESENZA DI PIU' DISTRIBUZIONI STAZIONARIE IMPLICA CHE LA CATENA E' RIDUCIBILE. POICHE' IL LIMITE DI SUCCESSIONE P^n DIPENDEREBBE DALLA RIGA, NON ESISTE UNA DISTRIBUZIONE LIMITE UNICA CHE VALGA PER OGNI STATO INIZIALE.