# ITCT Homework 2

- R07944007 網媒碩一 林良翰

## 1. [Prob. 3.1.]

### a. Markov's Inequality

- By the definition of probability
  $E\left[X\right] = \int_{-\infty}^{\infty} xP\left(x\right)dx$
- Since r.v. $X$ is non-negative
  $E\left[X\right] = \int_{0}^{\infty} xP\left(x\right)dx$
- By the property of integrals
  $E\left[X\right] = \int_{0}^{t} xP\left(x\right)dx + \int_{t}^{\infty} xP\left(x\right)dx \geq \int_{t}^{\infty} xP\left(x\right)dx \geq \int_{t}^{\infty} tP\left(x\right)dx = tP\left(X \geq t\right)$
  Thus
  $P\left(X \geq t\right) \leq \frac{E[X]}{t}$

### b. Chebyshev's Inequality

- By the definition of mean and variance
  - Mean : $\mu = E\left[Y\right]$
  - Variance : $\sigma^2 = E\left[\left(Y - E\left[Y\right]\right)^2\right] = E\left[\left(Y-\mu\right)^2\right] = E\left[X\right]$
- By Markov's enequality
  $P\left(X \geq \epsilon^2\right) \leq \frac{E[X]}{\epsilon^2}$
  Thus
  $P\left(\left|Y-\mu\right| \geq \epsilon\right) = P\left(\left(Y-\mu\right)^2 \geq \epsilon^2\right) \leq \frac{E\left[(Y-\mu)^2\right]}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}$

### c. Weak Law of Large Numbers

- Chebyshev's inequality
  A r.v. $X$ with its mean $E\left(X\right) = \mu$ and variance $Var\left(X\right) = \sigma^2$, for any $\epsilon \geq 0$
  $P\left(\left|X-\mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}$
  which is same as
  $P\left(\left|X - E\left(X\right)\right| \geq \epsilon\right) \leq \frac{Var(X)}{\epsilon^2}$
- Find $E\left(\bar{Z}_n\right)$ and $Var\left(\bar{Z}_n\right)$
  - $E\left(\bar{Z}_n\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} Z_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(Z_i\right) = \frac{1}{n}\left(n\mu\right) = \mu$
  - $Var\left(\bar{Z}_n\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right) = \frac{1}{n^2}Var\left(\sum_{i=1}^{n} Z_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left(Z_i\right) = \frac{1}{n^2}\left(n\sigma^2\right) = \frac{\sigma^2}{n}$
- Thus, by Chebyshev's inequality and the mean, variance of $\bar{Z}_n$, for any $\epsilon \geq 0$
  $P\left(\left|\bar{Z}_n - E\left(\bar{Z}_n\right)\right| \geq \epsilon\right) \leq \frac{Var(\bar{Z}_n)}{\epsilon^2}$
  $\Rightarrow P\left(\left|\bar{Z}_n - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$

# 2. [Prob. 3.4.]

## a.

- By AEP and the property of typical set ${A_\epsilon}^{(n)}$ with respect to $P(X)$ which is the set of sequence $(X_1, X_2, \ldots, X_n) \in X^n$
  $2^{-n(H(X)+\epsilon)} \leq P(X_1, X_2, \ldots, X_n) \leq 2^{-n(H(X)-\epsilon)}$
- Derive $A^n$
  $$\begin{aligned} A^n &= \left\{x^n \in X^n : \left|-\frac{1}{n}\log P(x^n) - H(X)\right| \leq \epsilon\right\} \\ &= \left\{x^n \in X^n : -\epsilon \leq -\frac{1}{n}\log P(x^n) - H(X) \leq \epsilon\right\} \\ &= \left\{x^n \in X^n : 2^{-n(H(X)+\epsilon)} \leq P(x^n) \leq 2^{-n(H(X)-\epsilon)}\right\} \\ &= {A_\epsilon}^{(n)} \end{aligned}$$
  $A^n$ is just the same as typical set
  $P\{X^n \in A^n\} = P\left\{X^n \in {A_\epsilon}^{(n)}\right\} \to 1-\epsilon$, $\forall \epsilon > 0$ as $n \to \infty$
  Thus, $P\{X^n \in A^n\} \to 1$

## b.

- By the strong law of large numbers
  $P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu\right) = 1$
  which can be rewrite as
  $\forall \epsilon > 0$, $P\left(\left|\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right| \leq \epsilon\right) \to 1$, as $n \to \infty$
- $B^n = \left\{x^n \in X^n : \left|\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right| \leq \epsilon\right\}$
  $P\{X^n \in B^n\} \to 1$
- By $P\{X^n \in A^n\} \to 1$, there exists $\epsilon > 0$ and $N_1$ such that $P\{X^n \in A^n\} > 1 - \frac{\epsilon}{2}$ for all $n > N_1$
  By $P\{X^n \in B^n\} \to 1$, there exists $\epsilon > 0$ and $N_2$ such that $P\{X^n \in B^n\} > 1 - \frac{\epsilon}{2}$ for all $n > N_2$
- So, for all $n > N = \max(N_1, N_2)$ and for any $\epsilon > 0$
  $$\begin{aligned} P\{X^n \in A^n \cap B^n\} &= P\{X^n \in A^n\} + P\{X^n \in B^n\} - P\{X^n \in A^n \cup B^n\} \\ &> 1 - \frac{\epsilon}{2} + 1 - \frac{\epsilon}{2} - 1 \\ &= 1 - \epsilon \end{aligned}$$
  Thus
  $P\{X^n \in A^n \cap B^n\} \to 1 - \epsilon$, for all $n > N$ and for any $\epsilon > 0$

## c.

- By the property of probability
  $\sum_{x^n \in A^n \cap B^n} P(x^n) \leq 1$
- By the property of typical set, for $x^n \in A^n$
  $P(x^n) \geq 2^{-n(H(X)+\epsilon)}$
- Combine above two inequations

$1 \geq \sum_{x^n \in A^n \cap B^n} P(x^n) \geq \sum_{x^n \in A^n \cap B^n} 2^{-n(H(X)+\epsilon)} \geq |A^n \cap B^n| \, 2^{-n(H(X)+\epsilon)}$

Therefore

$|A^n \cap B^n| \leq 2^{n(H(X)+\epsilon)}$

## d.

- From b.

  $P\{X^n \in A^n \cap B^n\} \to 1-\epsilon$, for all $n > \max(N_1, N_2)$ and for any $\epsilon > 0$

  Thus, there exists a number $N$ such that $P\{X^n \in A^n \cap B^n\} \geq \frac{1}{2}$, for all $n > N$
- By the property of typical set, for $x^n \in A^n$

  $P(x^n) \leq 2^{-n(H(X)-\epsilon)}$
- Combine above two inequations

  $\frac{1}{2} \leq P\{X^n \in A^n \cap B^n\} = \sum_{x^n \in A^n \cap B^n} P(x^n) \leq \sum_{x^n \in A^n \cap B^n} 2^{-n(H(X)-\epsilon)} = |A^n \cap B^n| \, 2^{-n(H(X)-\epsilon)}$

  Therefore

  $|A^n \cap B^n| \geq \left(\frac{1}{2}\right) 2^{n(H(X)-\epsilon)}$

# 3. [Prob. 3.10.]

- Define the notataions

  $\log x = \log_2 x$

  $\ln x = \log_e x$
- The volume $V_n = \prod_{i=1}^{n} X_i$

  We can derive

  $\log (V_n)^{\frac{1}{n}} = \frac{1}{n} \log (V_n) = \frac{1}{n} \sum_{i=1}^{n} \log (X_i)$
- By the strong law of large numbers, and $X$ is uniform over $[0, 1]$

  $\frac{1}{n} \sum_{i=1}^{n} \log (X_i) \to E[\log(X)] = \int_0^1 \log(x)\, dx = -\frac{1}{\ln 2}$
- Finally we can derive $\lim_{n\to\infty} (V_n)^{\frac{1}{n}}$

  $\lim_{n\to\infty} (V_n)^{\frac{1}{n}} = e^{\lim_{n\to\infty} \ln (V_n)^{\frac{1}{n}}} = e^{\ln 2 \left(\lim_{n\to\infty} \log (V_n)^{\frac{1}{n}}\right)} = e^{\ln 2 \left(-\frac{1}{\ln 2}\right)} = \frac{1}{e}$
- $(E[V_n])^{\frac{1}{n}} = \left(E\left[\prod_{i=1}^{n} X_i\right]\right)^{\frac{1}{n}} = \left(\prod_{i=1}^{n} E[X_i]\right)^{\frac{1}{n}} = \left(\left(\frac{1}{2}\right)^n\right)^{\frac{1}{n}} = \frac{1}{2} \neq \frac{1}{e}$

  Thus, the expected edge length does not capture the idea of the volume of the box.

# 4. [Prob. 3.11.]

## a.

- By definition and the property of probability

  $P(A) > 1 - \epsilon_1$

  $P(B) > 1 - \epsilon_2$

  $P(A \cup B) \leq 1 \Rightarrow -P(A \cup B) \geq -1$
- We can obtain the probablility of the intersection of $A$ and $B$

$P(A \cap B) = P(A) + P(B) - P(A \cup B) \geq 1 - \epsilon_1 + 1 - \epsilon_2 - 1 = 1 - \epsilon_1 - \epsilon_2$

## b.

- (a) from a.

  $1 - \delta - \epsilon \leq P\left({A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}\right)$

- (b) by the definition of probability of a set

  $P\left({A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}\right) = \sum_{x^n \in {A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}} P(x^n)$

- (c) by the property of typical set that $x^n \in {A_\epsilon}^{(n)}$

  $P(x^n)$ has upper bound $2^{-n(H-\epsilon)}$

  $\sum_{x^n \in {A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}} P(x^n) \leq \sum_{x^n \in {A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}} 2^{-n(H-\epsilon)}$

- (d) by the definition of the cardinality of a set

  $\sum_{x^n \in {A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}} 2^{-n(H-\epsilon)} = \left|{A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}\right| 2^{-n(H-\epsilon)}$

- (e) by the property that ${A_\epsilon}^{(n)} \cap {B_\delta}^{(n)} \subseteq {B_\delta}^{(n)}$

  $\left|{A_\epsilon}^{(n)} \cap {B_\delta}^{(n)}\right| 2^{-n(H-\epsilon)} \leq \left|{B_\delta}^{(n)}\right| 2^{-n(H-\epsilon)}$

## c.

- We apply similar steps in b.

  $$\begin{aligned} 1 - \delta &\leq P\left({B_\delta}^{(n)}\right) \\ &= \sum_{x^n \in {B_\delta}^{(n)}} P(x^n) \\ &\leq \sum_{x^n \in {B_\delta}^{(n)}} 2^{-n(H-\delta)} \\ &= \left|{B_\delta}^{(n)}\right| 2^{-n(H-\delta)} \end{aligned}$$

- Let $\delta' = \delta - \frac{1}{n}\log(1-\delta) > 0$

  Finally we can get

  $\frac{1}{n}\log\left|{B_\delta}^{(n)}\right| \geq H - \delta'$

# 5. [Prob. 4.6.]

## a.

- By the chain rule for entropy

  $$\begin{aligned} \frac{1}{n}H(X_1, X_2, \ldots, X_n) &= \frac{1}{n}\sum_{i=1}^{n} H(X_i | X_{i-1}, \ldots, X_1) \\ &= \frac{1}{n}\left(H(X_n | X_{n-1}, \ldots, X_1) + \sum_{i=1}^{n-1} H(X_i | X_{i-1}, \ldots, X_1)\right) \\ &= \frac{1}{n}(H(X_n | X_{n-1}, \ldots, X_1) + H(X_1, X_2, \ldots, X_{n-1})) \end{aligned}$$

- From stationarity, for all $1 \leq i \leq n$

  $H(X_n | X_{n-1}, \ldots, X_1) \leq H(X_i | X_{i-1}, \ldots, X_1)$

  We can derive

$H\left(X_n|X_{n-1},\ldots,X_1\right) \le \frac{1}{n-1}\sum_{i=1}^{n-1} H\left(X_i|X_{i-1},\ldots,X_1\right) = \frac{1}{n-1}H\left(X_1,X_2,\ldots,X_{n-1}\right)$

- Finally we can get

  $\frac{1}{n}H\left(X_1,X_2,\ldots,X_n\right) \le \frac{1}{n}\left(\frac{1}{n-1}H\left(X_1,X_2,\ldots,X_{n-1}\right) + H\left(X_1,X_2,\ldots,X_{n-1}\right)\right) = \frac{1}{n-1}H\left(X_1,X_2,\ldots,X_{n-1}\right)$

## b.

- By the stationarity, for all $1 \le i \le n$

  $H\left(X_n|X_{n-1},\ldots,X_1\right) \le H\left(X_i|X_{i-1},\ldots,X_1\right)$

- Thus we can derive

$$\begin{aligned} H\left(X_n|X_{n-1},\ldots,X_1\right) &= \frac{1}{n}\sum_{i=1}^{n} H\left(X_n|X_{n-1},\ldots,X_1\right) \\ &\le \frac{1}{n}\sum_{i=1}^{n} H\left(X_i|X_{i-1},\ldots,X_1\right) \\ &= \frac{1}{n}H\left(X_1,X_2,\ldots,X_n\right) \end{aligned}$$

# 6. [Prob. 4.8.]

- By the definition of entropy and $p_1 + p_2 = 1$

  $H(X) = -p_1\log p_1 - p_2\log p_2 = -p_1\log p_1 - (1-p_1)\log(1-p_1)$

- Define r.v. $T = \{t_1, t_2\}$ is the symbol duration of r.v. $X$

  $E[T] = p_1t_1 + p_2t_2 = p_1 + 2p_2 = 2 - p_1$

- Define $f(p_1)$ is the source entropy per unit time controled by $p_1$

  $f(p_1) = \frac{H(X)}{E[T]} = \frac{-p_1\log p_1 - (1-p_1)\log(1-p_1)}{2-p_1}$

- Since $f(0) = f(1) = 0$, the maximum point occurs between $0 < p_1 < 1$, thus we take 1st derivative of $f$ and set it to 0 to find optimal $p_1^\star$ such that $f(p_1^\star)$ has maximum

  $\frac{d}{dp_1}\left(\frac{-p_1\log p_1 - (1-p_1)\log(1-p_1)}{2-p_1}\right) = \frac{1}{(2-p_1)^2}\left[\ln(1-p_1) - 2\ln p_1\right] = 0$

  We can get

  $1 - p_1 = (p_1)^2 \Rightarrow p_1^\star = -\frac{1}{2} \pm \frac{\sqrt{5}}{2}$

  Since $0 < p_1 < 1$

  $p_1^\star = -\frac{1}{2} + \frac{\sqrt{5}}{2}$

- Organize the equation

$$\begin{aligned} f(p_1) &= \frac{-p_1\log p_1 - (1-p_1)\log(1-p_1)}{2-p_1} \\ &= \frac{-p_1\log p_1 - (1-p_1)\log(1-p_1)}{1+(1-p_1)} \\ &= \frac{-p_1\log p_1 - (p_1)^2\log(p_1)^2}{1+(p_1)^2} \\ &= \frac{-\left(1-(p_1)^2\right)\log p_1 - 2(p_1)^2\log p_1}{1+(p_1)^2} \\ &= \frac{-\left(1+(p_1)^2\right)\log p_1}{1+(p_1)^2} \\ &= -\log p_1 \end{aligned}$$

  Finally we can get

$-\log p_1^* = -\log\left(-\frac{1}{2}+\frac{\sqrt{5}}{2}\right) \approx 0.6942$

# 7. [Prob. 4.12.]

## a.

- The dog's walk is 2nd order Markov, because we need the previous 2 position to know the previous position and direction.
- By the chain rule of entropy

  $H\left(X_0, X_1, \ldots, X_n\right) = \sum_{i=0}^{n} H\left(X_i | X_{i-1}, \ldots, X_0\right) = H\left(X_0\right) + H\left(X_1 | X_0\right) + \sum_{i=2}^{n} H\left(X_i | X_{i-1}, \ldots, X_0\right)$
- The initial position is deterministic $\Rightarrow H\left(X_0\right) = 0$

  The first move is equally likely to positive or negative $\Rightarrow H\left(X_1 | X_0\right) = H\left(\frac{1}{2}, \frac{1}{2}\right) = 1$

  The rest of the move follow the rule defined by the problem

  $\sum_{i=2}^{n} H\left(X_i | X_{i-1}, \ldots, X_0\right) = \sum_{i=2}^{n} H\left(\frac{1}{10}, \frac{9}{10}\right) = (n-1) H\left(\frac{1}{10}, \frac{9}{10}\right)$
- Thus we have

  $H\left(X_0, X_1, \ldots, X_n\right) = 1 + (n-1) H\left(\frac{1}{10}, \frac{9}{10}\right)$

## b.

- The entropy rate

  $\frac{H(X_0, X_1, \ldots, X_n)}{n+1} = \frac{1+(n-1) H\left(\frac{1}{10}, \frac{9}{10}\right)}{n+1} \rightarrow H\left(\frac{1}{10}, \frac{9}{10}\right)$ as $n \rightarrow \infty$

## c. ?

- Let r.v. $S$ be the number of steps taken between reversals, we have

  $E(S) = \sum_{s=1}^{\infty} s\left(\frac{9}{10}\right)^{s-1} \frac{1}{10} = 10$
- Thus the expected number of steps the dog takes before reversing direction is

  $10 + 1 = 11$

# 8. [Prob. 4.19.]

- Define r.v. $V_n \in \{1,2,3,4,5\}$ is the vertex position at time $n$
- Define r.v. $X_n = \begin{bmatrix} P\left(V_n = 1\right) \\ P\left(V_n = 2\right) \\ P\left(V_n = 3\right) \\ P\left(V_n = 4\right) \\ P\left(V_n = 5\right) \end{bmatrix}$ is the vertex position distribution at time $n$
- The state transition matrix $p$

  $P_{ij} = P\left(V_{n+1} = j | V_n = i\right) = \begin{bmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix}$

## a.

- The stationary distribution of vertex position $X_n^*$

  $X_n^* = \lim_{n\to\infty} X_n = \begin{bmatrix} \frac{3}{16} & \frac{3}{16} & \frac{3}{16} & \frac{3}{16} & \frac{4}{16} \end{bmatrix}^T$ such that $X_n^* = P_{ij} \cdot X_n^*$

## b.

- By the definition of entropy rate

  $H(X) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)$

- Since $X_n$ is a stationary process and it's also a Markov chain

  $H(X) = H'(X) = \lim_{n\to\infty} H(X_n|X_{n-1}, \ldots, X_0) = H(X_1|X_0)$

  $H(X_1|X_0) = -\sum_{i=1}^{5} \left( P(V_0 = i) \sum_{j=1}^{5} P(V_1 = j|V_0 = i) \log P(V_1 = j|V_0 = i) \right) = 4 \cdot \frac{3}{16} \log 3 + \frac{4}{16} \log 4$

## c.

- By the definition of mutual information

$$\begin{aligned} I(X_{n+1}; X_n) &= H(X_{n+1}) - H(X_{n+1}|X_n) \\ &= H(X_{n+1}) - H(X_1|X_0) \\ &= \left[4 \cdot \frac{3}{16} \log \frac{16}{3} + \frac{4}{16} \log \frac{16}{4}\right] - \left[4 \cdot \frac{3}{16} \log 3 + \frac{4}{16} \log 4\right] \\ &= \frac{3}{4} \log \frac{16}{9} \end{aligned}$$

# 9.

- Define the notation

  $X \in \{a, b, c\}$

  $Y \in \{a', b', c'\}$

  and we have the following equation

  $\bar{P} + 2\bar{Q} = 1$

- By the definition of channel capacity

  $C = \max_{P_X} I(X; Y)$

- Find the formula of mutual information

$$\begin{aligned} I(X; Y) &= H(X) - H(Y|X) \\ &= -\bar{P} \ln \bar{P} - 2\bar{Q} \ln \bar{Q} - \left[ -\sum_{x\in X} \sum_{y\in Y} P(X = x, Y = y) \ln P(X = x|Y = y) \right] \\ &= -\bar{P} \ln \bar{P} - 2\bar{Q} \ln \bar{Q} + 2\bar{Q} q \ln q + 2\bar{Q}(1 - q) \ln(1 - q) \\ &= -\left(1 - 2\bar{Q}\right) \ln\left(1 - 2\bar{Q}\right) - 2\bar{Q} \ln \bar{Q} + 2\bar{Q} q \ln q + 2\bar{Q}(1 - q) \ln(1 - q) \end{aligned}$$

- To find the maximum value, we take 1st derivative of mutual information with $\bar{Q}$ and set it to 0

$$\begin{aligned} \frac{dI(X; Y)}{d\bar{Q}} &= \frac{d\left[-\left(1 - 2\bar{Q}\right) \ln\left(1 - 2\bar{Q}\right) - 2\bar{Q} \ln \bar{Q} + 2\bar{Q} q \ln q + 2\bar{Q}(1 - q) \ln(1 - q)\right]}{d\bar{Q}} \\ &= 2 \ln\left(1 - 2\bar{Q}\right) - 2 \ln \bar{Q} + 2\left[q \ln q + (1 - q) \ln(1 - q)\right] = 0 \end{aligned}$$

  Let $\alpha = q^q (1 - q)^{(1-q)}$

  The optimal $\bar{Q}^* = \frac{\alpha}{1+2\alpha}$

- Set $\bar{Q} = \bar{Q}^*$

$$
\begin{aligned}
I_{\bar{Q}^*}(X;Y) &= -\left(1-2\bar{Q}^*\right)\ln\left(1-2\bar{Q}^*\right) - 2\bar{Q}^*\ln\bar{Q}^* + 2\bar{Q}^* q\ln q + 2\bar{Q}^*(1-q)\ln(1-q) \\
&= -\left(1-2\frac{\alpha}{1+2\alpha}\right)\ln\left(1-2\frac{\alpha}{1+2\alpha}\right) - 2\frac{\alpha}{1+2\alpha}\ln\frac{\alpha}{1+2\alpha} + 2\frac{\alpha}{1+2\alpha}q\ln q + 2\frac{\alpha}{1+2\alpha}(1-q)\ln(1-q) \\
&= \ln(1+2\alpha) \\
&= \ln\left(1+2q^q(1-q)^{(1-q)}\right)
\end{aligned}
$$

## 10.

### a.

- $P(y_0) = P(x_0, y_0) + P(x_1, y_0) = \frac{1}{2}\frac{99}{100} + \frac{1}{2}\frac{1}{100} = \frac{1}{2}$
- $P(y_1) = P(x_0, y_1) + P(x_1, y_1) = \frac{1}{2}\frac{1}{100} + \frac{1}{2}\frac{99}{100} = \frac{1}{2}$

### b.

- $H(Y) = \frac{1}{2}\log 2 + \frac{1}{2}\log 2 = 1$

### c.

- By the definition of mutual information

$$
\begin{aligned}
I(X;Y) &= H(X) - H(Y|X) \\
&= 1 - \left[\frac{1}{2}\frac{99}{100}\log\frac{100}{99} + \frac{1}{2}\frac{1}{100}\log\frac{100}{1} + \frac{1}{2}\frac{99}{100}\log\frac{100}{99} + \frac{1}{2}\frac{1}{100}\log\frac{100}{1}\right] \\
&= 1 - \frac{99}{100}\log\frac{100}{99} - \frac{1}{100}\log\frac{100}{1} \\
&\approx 0.9192
\end{aligned}
$$