Challenge

Q1) $H(X) = -Pr(x=red)\log_2 Pr(x=red) - Pr(x=black)\log_2 Pr(x=black)$

(a) For maximum entropy, $H'(X)=0$

we know that,

$Pr(x=red) + Pr(x=black) = 1$

$Pr(x=black) = 1 - Pr(x=red)$

$H(x) = -P(x=red)\log_2 Pr(x=red) - (1-Pr(x=red))\log_2(1-Pr(x=red))$

$H'(x) = -\log P - \frac{P}{P} - \left(-\log(1-P) - \frac{(1-P)}{(1-P)}\right) = 0$

$\Rightarrow -\log P - \cancel{1} + \log(1-P) + \cancel{1} = 0$

$\Rightarrow \log\left(\frac{1-P}{P}\right) = 0$

$1-P = P \Rightarrow 2P = 1 \Rightarrow P = 1/2$

$\boxed{\text{To maximize entropy, buy 5 red socks and 5 black socks}}$

(b) $\underset{P}{\text{maximize}}\ f(P) := -P\log_2(P) - (1-P)\log_2(1-P)$

$f'(P) = -\log P - \frac{P}{P} - \left[-\log(1-P) + \frac{1-P}{1-P}(-1)\right] = 0$

$\Rightarrow -\log P - \cancel{1} + \log(1-P) + \cancel{1} = 0$

$$\Rightarrow \log\left(\frac{c-p}{p}\right) = 0$$

$$\Rightarrow c-p = p \quad \Rightarrow c = 2p \qquad p = c/2$$

$$\boxed{\therefore\ p = c/2 \quad \text{for optimum value}}$$

(c) Assume we have 3 different colours,

$P(\text{colour } 1) = x$

$P(\text{colour } 2) = y$

$P(\text{colour } 3) = z = 1-x-y$

$$H(x,y) = -x\log x - y\log y - (1-x-y)\log(1-x-y)$$

$$\frac{\partial H(x,y)}{\partial(x)} = 0, \quad \frac{\partial H(x,y)}{\partial(y)} = 0 \quad \Big\}\ \text{For maximum entropy}$$

$$\frac{\partial H(x,y)}{\partial(x)} = 0, \text{ we get } \quad 2x + y = 1 \quad \text{--- (1)}$$

$$\frac{\partial H(x,y)}{\partial(y)} = 0, \text{ we get } \quad 2y + x = 1 \quad \text{--- (2)}$$

Solving (1) & (2), we get $x = \frac{1}{3}$, $y = \frac{1}{3}$, $z = \frac{1}{3}$

we can extend this to n-colours,

where $\boxed{P(i) = \frac{1}{n} \quad (i = 1 \dots n)}$

Part (a) of the question is when $n = 2$

Here, $n = 100$

$$P(i) = \frac{1}{100}$$

Total socks $= 10{,}000$

$\therefore$ we need to buy $\frac{10{,}000}{100} = 100$ socks

of each color.

Q2) $Pr(Y_1=1) = P$,

$Pr(Y_i=1 \mid Y_{i-1}=1) = c+P-cP$

$Pr(Y_i=-1 \mid Y_{i-1}=-1) = cP-P+1$

(a) We will prove this with induction.

for $k=1$:- $Pr(Y_1=1) = P \rightarrow$ given

for $k=n-1$:- $Pr(Y_{n-1}=1) = P \rightarrow$ Assumption

for $k=n$:- $Pr(Y_n=1) = Pr(Y_n=1 \mid Y_{n-1}=1)\, Pr(Y_{n-1}=1)$
$+ Pr(Y_n=1 \mid Y_{n-1}=-1)\, Pr(Y_{n-1}=-1)$

$$= (c+P-cP)(P) + (1-(cP-P+1))(1-P)$$

$$= (c+P-cP)(P) + (P-cP)(1-P)$$

$$= P$$

$$\boxed{\therefore\ Pr(Y_i=1) = P \quad \text{for all } i}$$

(b) $$\mathrm{corr}(U,V) = \frac{E[UV] - E[U]\,E[V]}{\sqrt{\mathrm{Var}(U)\ \mathrm{Var}(V)}}$$

(y_i, y_{i-1})	$y_i * y_{i-1}$	$P(y_i, y_{i-1}) = P(y_i \mid y_{i-1})\, P(y_{i-1})$
$(1, 1)$	1	$(1 + p - cp)\,(p)$
$(1, -1)$	-1	$(p - cp)\,(1-p)$
$(-1, 1)$	-1	$(1 - (-p + cp)\,(p)$
$(-1, -1)$	1	$c\,(cp - p + 1)\,(1-p)$

$$E[uv] = \sum_{i=1}^{n} (u_i v_i)\, Pr(u_i, v_i)$$

$$= 4p^2 - 4p^2 c \quad + 4pc + 1 - 4p$$

$$E[v] = 1 \cdot p - 1\,(1-p) = 2p - 1$$

$$E[u] = E[v] = 2p - 1$$

$$E[uv] - E[u]\,E[v] = 4pc\,(1-p)$$

$$E[v^2] = 1^2\, p + (-1)^2 (1-p) = p + 1 - p = 1$$

$$var(u) = E[u^2] - E(u)^2 = 1 - (2p-1)^2 = 4p(1-p)$$

$$var(u) = var(v) = 4p(1-p)$$

$$\text{corr}(u,v) = \frac{4pc\,(1-p)}{\sqrt{4p(1-p) * 4p(1-p)}} = c$$

$$\boxed{\therefore\ \text{corr}(u,v) = c}$$

c) For, $m=3$, majority votes (atleast 2 1's) are possible in the following scenario

Y_1	Y_2	Y_3
1	1	1
1	1	-1
1	-1	1
-1	1	1

$Pr(Y_i = 1 \mid Y_{i-1} = 1) = P_{11}$ $\qquad$ $Pr(Y_i = -1 \mid Y_{i-1} = -1) = P_{00}$

$Pr(Y_i = -1 \mid Y_{i-1} = 1) = 1 - P_{11}$ $\qquad$ $Pr(Y_i = 1 \mid Y_{i-1} = -1) = 1 - P_{00}$

$= p * P_{11} * P_{11}$

$+ p * P_{11} * (1-P_{11})$

$+ p * (1-P_{11}) * (1-P_{00})$

$$+ (1-p) * (1-p_{00}) * p_{11}$$

$$= p * \cancel{(p_{11})^2} + p \cancel{* p_{11}} - p * \cancel{(p_{11})^2}$$

$$+ p - p p_{00} - \cancel{p * p_{11}} + p * p_{11} * p_{00}$$

$$+ p_{11} - p_{00} p_{11} - p p_{11} + p p_{00} p_{11}$$

$$= p - p p_{00} + 2(p p_{00} p_{11}) + p_{11} - p_{00} p_{11}$$

$$- p p_{11}$$

$$= p + p_{11} - p p_{11} - p_{00} p_{11} - p p_{00} + 2(p p_{00} p_{11})$$