

1 Chapter 1

1.2. Show that the error probability is reduced by the use of R_3 by computing the error probability of this code for a binary symmetric channel with noise level f .

Solution. The probability that an error occurred in a regular transmission was $f \in \{0...1\}$. When using the redundancy code R_3 that probability dropped from f to $f^3 + \binom{3}{2}f^2(1-f)$

2 Chapter 2

2.2 Are the random variables X and Y in the joint ensemble of figure 2.2 independent?

Solution. I am assuming this is asking if the $P(q|u)$ has any relation to $P(z|u)$. To which I say no and if you add up all the rows you will not get a uniform value.

2.4 An urn contains K balls, of which B are black and $W = K - B$ are white. Fred draws a ball at random from the urn and replaces it, N times.

1. What is the probability distribution of the number of times a black ball is drawn, nB ?
2. What is the expectation of nB ?
3. What is the variance of nB ?
4. What is the standard deviation of nB ?
5. Give numerical answers for the cases $N = 400$, when $B = 2$ and $K = 10$.
6. Give numerical answers for the cases $N = 5$, when $B = 2$ and $K = 10$.

Solution.

1. The probability distribution of nB is $\binom{N}{i} \left(\frac{B}{K}\right)^i \left(\frac{K-B}{K}\right)^{N-i}$
2. The expectation nB is

$$\begin{aligned} E[nB] &= N((1/K) * B) \\ &= (NB)/K \end{aligned}$$

3. The variance of nB is $(\sum_{i=1}^n i^2 P(Nb = i)) - E[nB]^2$ Where e is our expected value computed above.
4. The standard deviation of nB is the square root of its variance.

5. (a)

$$\begin{aligned}
 E[nB] &= N((1/K) * B) \\
 &= 400((1/10)2) \\
 &= 400((0.1)2) \\
 &= 400(0.2) \\
 &= 80
 \end{aligned}$$

(b)

$$\begin{aligned}
 var[nB] &= \sum_{i=1}^n (ie)^2 \\
 &= \sum_{i=1}^{400} (i80)^2 \\
 &= 6416000
 \end{aligned}$$

(c) $\sqrt{6416000} = 2532.98243184$

6. (a)

$$\begin{aligned}
 E[nB] &= N((1/K) * B) \\
 &= 5((1/10)2) \\
 &= 5((0.1)2) \\
 &= 5(0.2) \\
 &= 1
 \end{aligned}$$

(b)

$$\begin{aligned}
 var[nB] &= \sum_{i=1}^n (ie)^2 \\
 &= \sum_{i=1}^5 (i1)^2 \\
 &= 15
 \end{aligned}$$

(c) $\sqrt{15} = 3.87298334621$

2.10 Urn A contains three balls: one black, and two white; urn B contains three balls: two black, and one white. One of the urns is selected at random and one ball is drawn. The ball is black. What is the probability that the selected urn is urn A?

Solution. First the probability that we draw a black ball from urn A is $P(b|A) = 0.33333$ and the probability of drawing a black ball from urn B is $P(b|B) = 0.66666$. So if we draw a black ball it had a 0.33333 coming from urn A.

2.11 Urn A contains five balls: one black, two white, one green and one pink; urn B contains five hundred balls: two hundred black, one hundred white, 50 yellow, 40 cyan, 30 sienna, 25 green, 25 silver, 20 gold, and 10 purple. [One fifth of A's balls are black; two-fifths of B's are black.] One of the urns is selected at random and one ball is drawn. The ball is black. What is the probability that the urn is urn A?

Solution. First the probability that we draw a black ball from urn A is $P(b|A) = 0.2$ and the probability of drawing a black ball from urn B is $P(b|B) = 0.4$. So if we draw a black ball it had a 0.3333 coming from urn A.

2.26 Prove that the relative entropy (equation (2.45)) satisfies $DKL(P||Q) \geq 0$ (Gibbs' inequality) with equality only if $P = Q$.

Solution. If P and Q are the same then every probability is divided by itself giving you zero. If P and Q differ at all then you'll end up with two equations that are not the same p_1/q_1 and p_2/q_2

3 Chapter 8

8.7 Consider the ensemble XYZ in which $A_X = A_Y = A_Z = \{0, 1\}$, x and y are independent with $P_X = \{p, 1 - p\}$ and $P_Y = \{q, 1 - q\}$ and $z = (x + y) \bmod 2$

1. If $q = 1/2$, what is P_Z ? What is $I(Z; X)$?
2. For general p and q , what is P_Z ? What is $I(Z; X)$? Notice that this ensemble is related to the binary symmetric channel, with x = input, y = noise, and z = output.

Solution.

1. First let's define P_Z :

$$(a) P(z = 1) = P(x = 0, y = 1) + P(x = 1, y = 0) = 0.25 + 0.25 = 0.5$$

$$(b) P(z = 0) = P(x = 0, y = 0) + P(x = 1, y = 1) = 0.25 + 0.25 = 0.5$$

Now we can see that $H(Z) = 1$, $H(X) = 1$, and $H(X, Z) = 2$, even though X and Z are not independent variables Z is dependent on two independent variables one of which is X . $H(X, Z) = H(X, Y) = 2$. $I(Z, X) = H(X) - H(X|Z) = 0.5$.

2. The probability of Z will be the probability of X times Y . $I(X, Z)$ is the probability of X given Z .

8.1 Consider three independent random variables u, v, w with entropies $H(u), H(v), H(w)$. Let $X = (U, V)$ and $Y = (V, W)$. What is $H(X, Y)$? What is $H(X|Y)$? What is $I(X; Y)$?

Solution.

1. $I(X, Y) = H(v)$
2. $H(X, Y) = H(v) + H(u) + H(w)$
3. $H(X|Y) = H(u)$

8.4 Prove that the mutual information $I(X; Y) = H(X) - H(X|Y)$ satisfies $I(X; Y) = I(Y; X)$ and $I(X; Y) \geq 0$.

Solution.

4 Chapter 9

9.2 Now assume we observe $y = 0$. Compute the probability of $x = 1$ given $y = 0$.

Solution.

$$\begin{aligned}
 P(x = 1|y = 0) &= \frac{P(y = 0|x = 1)P(x = 1)}{\sum_{x'} P(y|x')P(x')} \\
 &= \frac{P(y = 0|x = 1)P(x = 1)}{P(y|x = 0)P(x = 0) + P(y|x = 1)P(x = 1)} \\
 &= \frac{0.15 * 0.1}{0.15 * 0.90 + 0.85 * 0.1} \\
 &= \frac{0.015}{0.22} \\
 &= 0.06818181818
 \end{aligned}$$

9.4 Alternatively, assume we observe $y = 0$ Compute $P(x = 1|y = 0)$.

Solution.

$$\begin{aligned}
 P(x = 1|y = 0) &= \frac{P(y = 0|x = 1)P(x = 1)}{\sum_{x'} P(y|x')P(x')} \\
 &= \frac{P(y = 0|x = 1)P(x = 1)}{P(y|x = 0)P(x = 0) + P(y|x = 1)P(x = 1)} \\
 &= \frac{0.15 * 0.1}{1 * 0.90 + 0.85 * 0.1} \\
 &= \frac{0.015}{0.985} \\
 &= 0.01522842639
 \end{aligned}$$

9.8 Compute the mutual information between X and Y for the Z channel with $f = 0.15$ when the input distribution is $P_X : p_0 = 0.5, p_1 = 0.5$.

Solution. 0.678788110264986

9.12 What is the capacity of the binary symmetric channel for general f ?

Solution. F is the noise of a channel, as F increases the capacity decreases until it hits zero at $f = 0.5$.

9.13 Show that the capacity of the binary erasure channel with $f = 0.15$ is $C_{BEC} = 0.85$. What is its capacity for general f ? Comment.

Solution. The capacity of the channel is $1 - f$ for all f .

9.14 Find the transition probability matrices Q for the extended channel, with $N = 2$, derived from the binary erasure channel having erasure probability 0.15. By selecting two columns of this transition probability matrix, we can define a rate-1/2 code for this channel with blocklength $N = 2$. What is the best choice of two columns? What is the decoding algorithm?

Solution. Assuming that we are using the optimal encoder such that 0 and 1 have equal probability and there is no educated guess about what an eraser could have been, we can simply assign partial erasures to a full codeword at random:

$$1? = 11$$

$$0? = 00$$

$$?0 = 10$$

$$?1 = 01$$

We can also declare $??$ to be a decoding error as there is no partially good choice for what it might be. Another option to make decoding easier and the code no less lossy is to assign every erasure to zero. The capacity of this channel is twice that of the regular binary erasure channel but only because we have increased the blocklength. Either way if we receive an erasure on a perfectly distributed input the entire information bit is lost.

13.7. Give a simple rule that distinguishes whether a binary vector is orthogonal to itself, as is each of the three vectors $[1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0]$, $[0111010]$, and $[1011001]$.

solution. If a vector has even weight.

13.24.

solution. If you can't figure anything out about your own hat and each hat is a 50/50 chance of being blue or red then no matter how many people are in the room every one but a single person should pass while one picks a random guess. Having multiple people guess only decreases your 50/50 chance of winning.

15.4. How can you use a coin to draw straws among 3 people?

solution. Flip the coin three times for each pair of people, choose the person with the most wins. Or flip the coin for two of the three then the winner flips with the remaining person.

14.5.

solution. The magician arranges the cards like so: Each card can be faced up and right side up, faced down and right side up, faced up and upside down, faced down and upside down. If we treat this as a based 4 number system we get 4^5 possible combinations and each permutation can count as a single number.

16.1. If one creates a 'random' path from A to B by flipping a fair coin at every junction where there is a choice of two directions, is the resulting path a uniform random sample from the set of all paths?

solution. No it is not because once you hit the far right then your only choice is down so instead count the number of paths going through this node down and the number going through to the left and weight the coin flip such that it is an even chance that any of the paths going through the node is selected.

16.2. Having run the forward and backward algorithms between points A and B on a grid, how can one draw one path from A to B uniformly at random?

solution. Count the number of paths going through this node down and the number going through to the left and weight the coin flip such that it is an even chance that any of the paths going through the node is selected.

16.4. Having run the forward and backward algorithms between points A and B on a grid, how can one draw one path from A to B uniformly at random?

solution. Yes this can be completed by message passing with these simple rules per node:

1. If your x is $x < x_2 - 1$ and your y is $y < y_2 - 1$ then take the value you received from your left and pass it to the right plus your own intensity.

2. If your x is $x = x_2 - 1$ and your y is $y < y_2 - 1$ then take the value you received from both above and from the left add them together and add you own intensity and pass it bellow you.
3. If your x is $x = x_2 - 1$ and your y is $y = y_2 - 1$ then take the value you received from both above and to the left add your own intensity and you have your result

For this to work for the whole image just set $x_1 = y_1 = 0$, and x_2 and y_1 to the end of the image.

25.2 Confirm that the sixteen codewords listed in table 1.14 are generated by the trellis shown in figure 25.1c.

Solution Up until time slice 3 each node splits into 2 nodes leaving us with 8 nodees and 8 possible paths/codewords, then it again splits exactly twice but overlaps such that while we still have 8 nodes we now have 16 codewords. The last 3 slices of time no nodes have any choice on where to go, leaving us with the 16 codewords.

25.4 Find the most probable codeword in the case where the normalized likelihood is (0.2, 0.2, 0.9, 0.2, 0.2, 0.2, 0.2). Also find or estimate the marginal posterior probability for each of the seven bits, and give the bit-by-bit decoding.

	t	Likelihood	P(t y)
Solution	0000000	0.0262144	0.300623853211009
	0001011	0.0004096	0.00469724770642202
	0010111	0.0036864	0.0422752293577982
	0011100	0.0147456	0.169100917431193
	0100110	0.0004096	0.00469724770642202
	0101101	0.0001024	0.00117431192660550
	0110001	0.0147456	0.169100917431193
	0111010	0.0036864	0.0422752293577982
	1000101	0.0004096	0.00469724770642202
	1001110	0.0001024	0.00117431192660550
	1010010	0.0147456	0.169100917431193
	1011001	0.0036864	0.0422752293577982
	1100011	0.0001024	0.00117431192660550
	1101000	0.0004096	0.00469724770642202
	1110100	0.0036864	0.0422752293577982
	1111111	0.0000576	0.000660550458715597

n	$P(y_n t_n = 1)$	$P(y_n t_n = 0)$	$P(t_n = 1 y)$	$P(t_n = 0 y)$
1	0.2	0.8	0.266055045871560	0.733944954128440
2	0.2	0.8	0.266055045871560	0.733944954128440
3	0.9	0.1	0.677064220183487	0.322935779816513
4	0.2	0.8	0.266055045871560	0.733944954128440
5	0.2	0.8	0.266055045871560	0.733944954128440
6	0.2	0.8	0.266055045871560	0.733944954128440
7	0.2	0.8	0.266055045871560	0.733944954128440

And if you attempt bit by bit decoding you will properly decode just the message digits and ignore the parity check digits resulting in a useless extension of just transmitting your message.

25.5 Show that for a node i whose time-coordinate is n , a_i is proportional to the joint probability that the codeword's path passed through node i and that the first n received symbols were $\{y_1, \dots, y_n\}$.

Solution $a_0 = 1$ and $a_i = \sum_{x=1}^i a_{x-1} * P(a_x|y)$

25.8 From Exercise 25.5 we get the probability at each node is the probability we passed it. Then by assigning the probabilities above to the probabilities $P(t_n|y)$, we end with $P(t_n = t|y) = \sum_{n=1}^N P(t_{n-1}|y) * P(t_n|y)$