Задачи по теории информации

1. Общие вопросы

1.1. Слабый закон больших чисел 1-1

Пусть $(X_1, \dots, X_N)$ – блок одинаковых независимых случайных величин со средним $m = E[X]$ и дисперсией
$$\sigma^2 = E[(X-m)^2].$$
Пусть
$$S_N = \frac{\sum_n x_n}{N}.$$
Покажите что
$$E[S_N] = m; \quad \sigma^2(S_N) = E[(S_N - m)^2] = \frac{\sigma^2}{N}.$$

1.2. Слабый закон больших чисел 5-1

Пусть $(X_1, \dots, X_N)$ – блок одинаковых независимых случайных величин со средним $m = E[X]$ и дисперсией
$$\sigma^2 = E[(X-m)^2].$$
Пусть
$$V_N = \frac{\sum_n x_n}{\sqrt{N}}.$$
Покажите что
$$E[V_N] = m\sqrt{N}; \quad \sigma^2(V_N) = E[(V_N - m\sqrt{N})^2] = \sigma^2.$$

1.3. К формуле Стирлинга 2-1

Получить нижнюю и верхнюю границы для факториала:
$$e^{7/8}[n^n e^{-n}\sqrt{n}] < n! < e[n^n e^{-n}\sqrt{n}].$$
Сопоставить ее с известной асимптотической оценкой Стирлинга: $n! \sim n^n e^{-n}\sqrt{2\pi n}$. $(e^{7/8} = 2.399 < \sqrt{2\pi} = 2.507 < e = 2.718)$.

Решение

Имеем:
$$\ln n! = \sum_{k=1}^{n} \ln k.$$
Площадь под кривой $\ln x$ легко вычисляется:
$$S_n = \int_1^n \ln x dx = \int_1^n d(x \ln x) - \int_1^n dx = \ln(n^n e^{-n}/e).$$

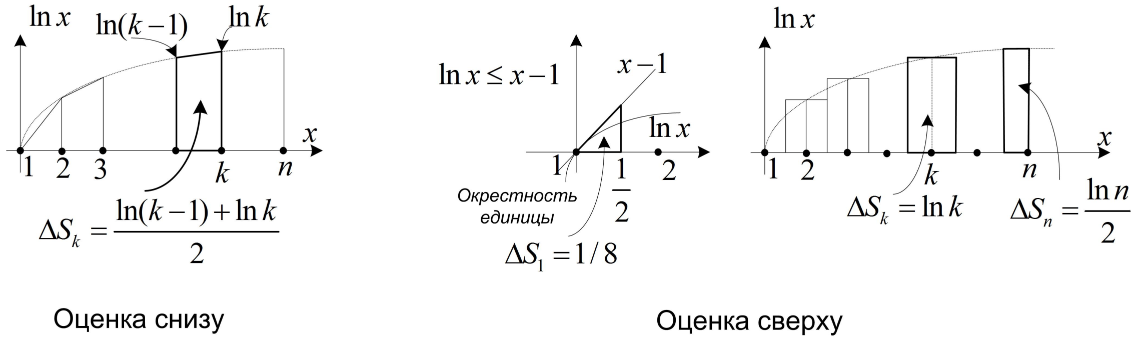

Рис. 1. Оценки площади под кривой логарифма

К успеху приводят оценки этой площади сверху и снизу через сумму для $\ln n!$, см. рисунок

Оценка снизу суммой площадей трапеций дает

$$S_n > \frac{\ln 1+\ln 2}{2}+\frac{\ln 2+\ln 3}{2}+\dots\frac{\ln(k-1)+\ln k}{2}+\frac{\ln k+\ln(k+1)}{2}+\dots+\frac{\ln(n-1)+\ln n}{2}=$$

$$=\ln n!-\frac{\ln n}{2}=\ln\frac{n!}{\sqrt{n}}.$$

Оценка сверху суммой площадей прямоугольников приводит к

$$S_n<\ln n!-\frac{\ln n}{2}+\frac{1}{8}=\ln\frac{e^{1/8}n!}{\sqrt{n}}.$$

Таким образом,

$$\ln\frac{e^{1/8}n!}{\sqrt{n}}<S_n=\ln(n^ne^{-n}/e)<\ln\frac{n!}{\sqrt{n}}$$

А это и есть требуемые границы для факториала.

1.4. Оценки биномиальных распределений 3-1

Показать, что

$$\binom{n}{k}\simeq 2^{h(\frac{k}{n})},$$

где $h(x)=-x\log x-(1-x)\log(1-x)$. Получить оценку для вероятности $P_n(k)$ выпадения k орлов в серии из n бросаний симметричной монеты:

$$P_n(k)=2^{-n}\binom{n}{k}\simeq 2^{-n(1-h(\frac{k}{n}))}$$

1.5. Оценки биномиальных распределений 6-1

Пусть монета несимметрична и вероятность выпадения орла составляет p. Получить оценку для вероятности $P_n(k)$ выпадения k орлов в серии из n испытаний

$$P_n(k)=\binom{n}{k}p^k(1-p)^{n-k}\simeq 2^{-nD(\frac{k}{n},p)},$$

где

$$D\Big(\frac{k}{n},p\Big)=\frac{k}{n}\log\frac{\frac{k}{n}}{p}+\frac{n-k}{n}\log\frac{\frac{n-k}{n}}{p}.$$

Показать, что наиболее вероятное значение $\frac{k}{n}$ равно p.

1.6. Неравенство Йенсена 4-1

Пусть $f(x)$ – выпуклая вниз функция, а $X = \{x\}$ – случайная величина со средним значением $E[x]$. Показать, что

$$E[f(x)] \geq f(E[x])$$

Решение

Значения выпуклой вниз функции лежат выше касательной, проведенной к ней в любой точке x_0:

$$f(x) \geq f(x_0) + \alpha(x - x_0).$$

Имеем

$$E(f(x)) \geq E[f(x_0) + \alpha(x - x_0)] = f(x_0) + \alpha(E[x] - x_0|).$$

Выбрав $x_0 = E[x]$, придем к

$$E(f(x)) \geq f(E(x)).$$

1.7. Применение неравенства Йенсена 7-1

Применив неравенство Йенсена, покажите, что центр масс системы из нанизанных на веревку шариков находится выше веревки, см. рисунок.

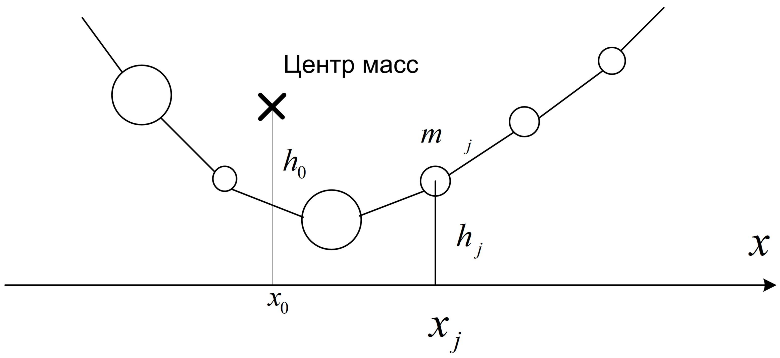

Рис. 2. Шарики на веревке

Решение

Функция $h(x)$, выражающая зависимость от x высоты веревки, выпукла вниз. Координаты центра масс – это

$$x_0 = \sum p_j x_j = E[x]; \quad h_0 = \sum p_j h_j = E[h(x)],$$

где $p_j = \frac{m_j}{M}$, $M = \sum m_j$. Но $E[h(x)] \geq h(E[x])$.

2. Свойства энтропийных функций

2.1. q-ичная энтропия 8-1

Найти формулу для энтропии $H(p)$ распределения на q-точках с $p_1 = 1 - p$ и $p_2 = \cdots = p_q = \frac{p}{q-1}$. При каком значении параметра p достигается максимум $H(p)$ и каково значение этого максимума ?

Решение

Определение энтропии дает:

$$H(p) = -p \log_2 p - (1-p)\log_2(1-p) + p\log_2(q-1) = h(p) + p\log_2(q-1).$$

Максимум достигается на равномерном распределении, когда $1 - p = \frac{p}{q-1}$, то есть $p = \frac{q-1}{q}$ и составляет $\log_2 q$, превышая значение $\log_2(q-1)$ при $p = 1$ на $\log_2 \frac{q}{(q-1)}$.

2.2. q-ичная энтропия 9-1

Найти формулу для энтропии $h_q(p) = -\sum_p p\log_q(p)$ по основанию q распределения на q-точках с $p_1 = 1 - p$ и $p_2 = \cdots = p_q = \frac{p}{q-1}$. При каком значении параметра p достигается максимум $h_q(p)$ и каково значение этого максимума ? Постройте набросок графика $h_q(p)$.

Решение

Переход к логарифмам по основанию q в предыдущей задаче дает

$$h_q(p) = -p\log_q p - (1-p)\log_q(1-p) + p\log_q(q-1)$$

с единичным максимумом в точке $p = \frac{q-1}{q}$ ($\log_q x = \frac{\log_2 x}{\log_2 q}$).

2.3. Границы для энтропии 10-1

Пусть случайная величина принимает M значений. Покажите, что

$$0 \le H(X) \le \log M.$$

При каких условиях нижняя и верхняя границы достигаются.

Решение

Энтропия

$$H(X) = -\sum_x p(x)\log p(x)$$

достигает минимального нулевого значения на распределении, принимающем значение 1 в одной точке и значения 0 в прочих. С другой стороны:

1. Применим неравенство логарифма: $\log x \le \log\, e(x-1)$.

$$H - \log M = \sum_x p(x)\log\frac{1}{p(x)M} \le \log\, e\sum_x p(x)\Big(\frac{1}{p(x)M} - 1\Big) = 0.$$

2. Применим неравенство Йенсена. Поскольку $\log x$ – выпуклая вверх функция,

$$H - \log M = \sum_x p_x \log\frac{1}{p_x M} = E\Big[\log\frac{1}{p_x M}\Big] \le \log E\Big[\frac{1}{p_x M}\Big] = \log 1 = 0.$$

3. Решим задачу максимизации $H(X)$ по распределениям $P = \{p(x)\}$ при условии $\sum p(x) = 1$. Приравняв к нулю компоненты градиента лагранжиана

$$L(P) = -\sum p(x)\log p(x) - \mu\sum p(x)$$

найдем

$$\frac{dL}{dp} = -\log p - 1 - \mu = 0; \quad \Rightarrow \quad p = e^{-1+\mu}.$$

То есть, максимум энтропии достигается на равномерном распределении.

4. Через информационную дивергенцию:

$$\log M - H = \sum_x p(x)\log\frac{p(x)}{1/M} = D(P||1/M) \ge 0.$$

2.4. Информационная дивергенция 11-1

Пусть $P = \{p(x)\}$ $Q = \{q(x)\}$ – два распределения вероятностей на одном и том же числе точек. Введем информационную дивергенцию между распределениями:

$$D(P||Q) = \sum_x p(x) \log \frac{p(x)}{q(x)}.$$

Покажите, что

$$D(P||Q) \geq 0.$$

При каком условии достигается равенство ?

Решение

1. Через неравенство логарифма

$$-D(P||Q) = \sum p(x) \log \frac{q(x)}{p(x)} \leq \log e \sum p(x) \Big(\frac{q(x)}{p(x)} - 1 \Big) = 0.$$

2. Через неравенство Йенсена

$$D(P||Q) = \sum p(x) \log \frac{p(x)}{q(x)} = E\Big[-\log \frac{q(x)}{p(x)}\Big] \geq -\log E\Big[\frac{q(x)}{p(x)}\Big] = 0.$$

2.5. К информационной дивергенции 12-1

Пусть $P = \{p(x)\}$ $Q = \{q(x)\}$ – два распределения вероятностей на одном и том же числе точек. Рассмотрим функционал

$$F(P||Q) = \sum_x p(x) \log \frac{1}{q(x)}.$$

На каком распределении P достигается максимум $F(P||Q)$? На каком распределении Q достигается минимум $F(P||Q)$? Вывести отсюда границу

$$D(P||Q) = \sum p(x) \log \frac{p(x)}{q(x)} = F(P||Q) - H(P) \geq 0$$

для информационной дивергенции.

Решение

Максимум по P достигается на распределении, равном единице в той точке x, в которой значение $\log \frac{1}{p(x)}$ максимально. Если таких точек несколько, то полную вероятность 1 можно распределить между ними как угодно. На значение $F(P||Q)$ это не влияет.

Чтобы найти экстремум по q, приравняем к нулю градиент лагранжиана

$$L(Q) = \sum p(x) \log \frac{1}{q(x)} + \mu \sum q(x).$$

Получим:

$$\frac{dL}{dq} = -\frac{p}{q} + \mu = 0.$$

Таким образом, минимум достигается при $\mu q(x) = p(x)$, то есть, с учетом нормировки $\sum q(x) = 1$, при $q(x) = p(x)$. Информационная дивергенция обращается в точке минимума $F(P||Q)$ в нуль.

2.6. Аддитивность энтропии 13-1

Пусть во множестве значений случайной величины $X = \{x_1, x_2, \dots, x_k, x_{k+1} \dots x_n\}$ с распределением $P = \{p_1, p_2, \dots, p_k, p_{k+1}, \dots p_n\}$ выделен агрегат $A = \{x_1, x_2, \dots x_k\}$ с полной вероятностью $p_A = \sum_{j=1}^{k} p_j$ и распределением $P_A = \{\frac{p_1}{p_A}, \dots, \frac{p_k}{p_A}\}$ и его дополнение – агрегат $\bar{A}$ с вероятностью $1 - p_A$ и распределением $P_{\bar{A}} = \{\frac{p_k}{1-p_A}, \dots, \frac{p_n}{1-p_A}\}$. Показать, что

$$H(X) = h(p_A) + p_A H(A) + (1 - p_A) H(\bar{A}).$$

Решение

$$-H(X) = \sum_{j=1}^{k} p_j \log p_j + \sum_{j=k}^{n} p_j \log p_j$$

$$= \sum_{j=1}^{k} p_A \frac{p_j}{p_A} \log p_A \frac{p_j}{p_A} + \sum_{j=k}^{n} (1 - p_A) \frac{p_j}{1 - p_A} \log(1 - p_A) \frac{p_j}{1 - p_A} =$$

$$= p_A \log p_A + p_A(-H(A)) + (1 - p_A) \log(1 - p_A) + (1 - p_A)(-H(\bar{A})).$$

Альтернативная интерпретация. Пусть U – индикатор множества A – случайная величина со значением 1 при $x \in A$ и 0 в противном случае. Тогда

$$H(X) = H(X, U) = H(U) + H(X/U),$$

где

$$H(U) = h(p_A)$$

а

$$H(X/U) = p_A H(A) + (1 - p_A) H(\bar{A}).$$

2.7. Укорочение случайной величины 14-1

Пусть $X_n = \{x_1, x_2, \dots, \dots x_n\}$ – случайная величина с распределением $P = (p_1 = p, p_2, \dots, p_n)$. Укоротим ее до случайной величины $X_{n-1} = \{x_2, \dots, \dots x_n\}$ с распределением $P = (\frac{p_2}{1-p}, \dots, \frac{p_n}{1-p})$. Показать, что

$$H(X_n) = h(p) + (1 - p) H(X_{n-1}).$$

Решение

Это частный случай предыдущей задачи при $A = \{x_1\}$.

2.8. Разбиение множества значений 15-1

Пусть множество значений случайной величины $X = \{x\}$ с распределением $P = \{p(x)\}$ разбито на M непересекающихся подмножеств X_m – агрегатов с полными вероятностями $P_m = \sum_{x \in X_m} p(x)$ и распределениями $P_m = \{p_m(x) = \frac{p(x)}{P_m}\}$. Показать, что

$$H(X) = H(P_1, P_2, \dots, P_M) + \sum_{j=1}^{M} P_m H(X_m).$$

Решение

Введем индикатор элемента разбиения – случайную величину U, принимающую разные значения на разных элементах. Ясно, что значение x волне определяет значение элемента U. Поэтому $H(U/X) = 0$ и $H(X,U) = H(X) + H(U/X) = H(X)$. С другой стороны,

$$H(X) = H(X,U) = H(U) + H(X/U) = H(P_1, P_2, \ldots, P_M) + \sum_{j=1}^{M} P_m H(X_m)$$

2.9. Оценки энтропийных функций 16-1

Пусть A, B, C – статистически независимые случайные величины с распределением $P_{ABC}(abc) = P_A(a)P_B(b)P_C(c)$ и энтропиями $H(A), H(B), H(C)$. Для случайных величин $X = (A, B)$ и $Y = (B, C)$ найти $I(X,Y)$ и $R(X,Y) = H(Y|X) + H(X|Y)$.

Решение

Независимость A, B, C дает

$$H(X) = H(AB) = H(A) + H(B),$$

$$H(Y) = H(BC) = H(B) + H(C),$$

Задание A, B, C вполне определяет значения X и Y. Поэтому

$$H(ABC) = H(XY) = H(A) + H(B) + H(C).$$

Далее:

$$H(X|Y) = H(XY) - H(Y) = H(A).$$

$$H(Y|X) = H(XY) - H(X) = H(C).$$

Потому

$$I(X,Y) = H(X) - H(X|Y) = H(B).$$

$$I(X,Y) = H(Y) - H(Y|X) = H(B).$$

$$R(X,Y) = H(X|Y) + H(Y|X) = H(A) + H(C).$$

$$I(X,Y) + R(X,Y) = H(XY).$$

2.10. Вычисление энтропийных функций 22-1

Пусть X, Y независимые случайные величины с равновероятными значениями 0 и 1. Введем случайные величины $S = (X+Y)\mathrm{mod}2$ и $M = XY$. Найти

$$H(SM), H(S), H(M), H(S|M), H(M|S), I(S,M), R(S,M) = H(S|M) + H(M|S)$$

Решение

На вероятностном пространстве из 4-х точек $(xy) = (00), (10), (01), (11)$ с вероятностными мерами $\frac{1}{4}$ заданы случайные функции со значениями

P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
XY	00	10	01	11
S	0	1	1	0
M	0	0	0	1

Случайная величина (SM) принимает три значения $(00), (10), (0,1)$ с вероятностями $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$. Поэтому

$$H(SM) = \frac{3}{2}.$$

Случайная величина S принимает значения $0, 1$ с вероятностями $\frac{1}{2}, \frac{1}{2}$. Так что

$$H(S) = 1.$$

Наконец, значения $0, 1$ величины M встречаются с вероятностями $\frac{3}{4}, \frac{1}{4}$. Такт что

$$H(M) = 2 - \frac{3}{4}\log 3.$$

Далее:

$$H(S|M) = P(M=0)H(S|M=0)+P(M=1)H(S|M=1) = P(M=0)H(S|M=0)+0 =$$

$$= \frac{3}{4}H(S|M=0).$$

При $M = 0$ S принимает значение 0 с вероятностью $\frac{1/4}{3/4} = 1/3$ и значение 1 с вероятностью $\frac{2/4}{3/4} = 2/3$. Поэтому $H(S|M=0) = \log 3 - 2/3$ и

$$H(S|M) = \frac{3}{4}\log 3 - \frac{1}{2}$$

. Наконец,

$$H(M|S) = P(S=0)H(M|S=0)+P(S=1)H(M|S=1) = P(S=0)H(M|S=0)+0 =$$

$$= \frac{1}{2}H(M|S=0).$$

Но $H(M|S=0) = 1$. Так что

$$H(M|S) = \frac{1}{2}$$

2.11. Вычисление энтропийных функций 17-1

Пусть X, Y независимые случайные величины с равновероятными значениями 0 и 1. Введем случайные величины $S = (X+Y)\text{mod}2$ и $M = XY$. Проверить что

$$H(SM) = H(S)+H(M|S) \quad H(SM) = H(M)+H(S|M) \quad H(SM) = I(S,M)+R(S,M),$$

где

$$R(S,M) = H(M|S) + H(S|M)$$

Решение

См. предыдущую задачу

2.12. Добавление случайной величины 18-1

Покажите что энтропия совместного распределения больше энтропии маргинального:

$$H(XY) \geq H(X); \quad H(XY|Z) \geq H(X|Z).$$

При каких условиях достигаются равенства.

Решение

Цепное правило дает

$$H(XY) = H(X) + H(Y|X) \geq H(X)$$

с равенством при $H(Y|X) = 0$, то есть когда X вполне определяет Y.

$$H(XY|Z)\sum_z P(Z)H(XY|Z=z) = \sum_z P(Z)(H(X|Z=z) + H(Y|XZ=z)) \geq$$

$$\geq \sum_z P(Z)H(X|Z=z) = H(X|Z).$$

2.13. Добавление условия 19-1

Покажите что добавление условия снижает энтропию:

$$H(X|Y) \leq H(X); \quad H(X|YZ) \leq H(X|Z).$$

При каких условиях достигаются равенства.

Решение

$$H(X|Y) = \sum_y P_Y(y)H(X|Y=y) = -\sum_y P_Y(y) \sum_x P(x|y)\log(P(x|y)] = -\sum_{xy} P(xy)\log P(x_y)$$

Так что

$$H(X) - H(X|Y) = \sum_{xy} P(xy)\log\frac{p(x|y)}{p(x)} = \sum_{xy} P(xy)\log\frac{p(x|y)p(y)}{p(x)p(y)} =$$

$$\sum_{xy} P(xy)\log\frac{p(xy)}{p(x)p(y)} = D(P(XY)||P(X)P(Y)) \geq 0$$

с равенством при $P(XY) = P(X)P(Y)$, то есть при независимости X и Y.

Далее

$$H(X|YZ) = \sum P_Z(z)H(X|Y, Z=z); \quad H(X|Z) = \sum P_Z(z)H(X|Z=z).$$

Неравенство в среднем $H(X|YZ) \leq H(X|Z)$ имеем место, поскольку оно справедливо в частности, при каждом z. Ясно , что

$$H(X|Z) - H(X|YZ) = \sum_{xyx} P(xyz)\log\frac{P(x|yz)}{P(x|y)} = \sum_{xyx} P(xyz)\log\frac{P(x|yz)P(yz)}{P(x|y)P(yz)} =$$

$$= \sum_{xyx} P(xyz)\log\frac{P(xyz)}{P(x|y)P(z)P(y|z)} = D(P(XYZ)||P(Z)P(Y|Z)P(X|Y) \geq 0$$

с равенством, когда величины Z, Y, X образуют цепь Маркова.

2.14. Добавление случайной величины 20-1

Покажите, что

$$H(Y,Z) - H(Z) \geq H(X,Y,Z) - H(X,Y).$$

(Добавление X повышает разность между энтропиями совместного и маргинального распределений)

Решение

Это просто завуалированное неравенство:

$$H(Y,Z) - H(Y) = H(Z|Y) \geq H(Z|XY) = H(X,Y,Z) - H(X,Y)$$

2.15. Мера взаимной случайности 21-1

Пусть $R(X,Y) = H(X|Y) + H(Y|X)$ – мера взаимной случайности между X и Y. Покажите, что

$$0 \le R(X,Y) \le H(X) + H(Y).$$

При каких условиях эти границы достигаются ? Проверьте выполнение неравенства треугольника:

$$R(X,Y) \le R(X,Z) + R(Z,Y).$$

Решение

Прежде всего, $H(X|Y) \ge 0$, $H(Y|X) \ge 0$ с равенствами, когда X и Y детерминированно связаны. Если же они независимы, то $H(X|Y) = H(X)$, $H(Y|X) = H(Y)$.

Далее

$$R(X,Y) = H(X|Y) + H(Y|X) \le H(XZ|Y) + H(YZ|X) =$$

$$= H(Z|Y) + H(X|YZ) + H(Z|X) + H(Y|XZ) \le$$

$$\le H(Z|Y) + H(X|Z) + H(Z|X) + H(Y|Z) = R(Y,Z) + R(X,Z).$$

2.16. К взаимной информации 23-1

Докажите эквивалентность представлений для взаимной информации

$$I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) =$$

$$= H(X) + H(Y) - H(X,Y) = H(X,Y) - R(X,Y).$$

Покажите, что $I(X,Y) \le H(X)$, $I(X,Y) \le H(Y)$. Когда эти границы достигаются.

Решение

Цепное правило дает

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y),$$

или

$$H(Y|X) = H(XY) - H(X); \quad H(X|Y) = H(XY) - H(Y).$$

Поэтому

$$I(X,Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y) = H(Y) + H(Y|X).$$

Далее,

$$H(X,Y) - R(X,Y) = H(X,Y) - (H(X|Y) + H(Y|X)) =$$

$=$

$$= (H(X,Y) - H(X|Y)) - H(Y|X) = H(Y) - H(Y|X) = I(X,Y).$$

2.17. К взаимной информации 24-1

Покажите, что совместная взаимна информация превышает маргинальную:

$$I(XY,Z) \ge I(X,Z),$$

$$I(XY,Z) \ge I(Y,Z).$$

При каких условиях достигаются равенства.

Решение

$$I(XY,Z) = H(XY) - H(XY|Z) = H(X) + H(Y|X) - H(X|Z) - H(Y|XZ) =$$

$$= I(X,Z) + (H(Y|X) - H(Y|XZ)) \geq I(X,Z)$$

поскольку $H(Y|X)-H(Y|XZ) \geq 0$. Равенство достигается когда $H(Y|X)-H(Y|XZ)$, то есть когда последовательность $Z \rightarrow X \rightarrow Y$ образует марковскую цепь : $(P(y|zx) = P(y|x))$.

2.18. Шифрование 25-1

Пусть открытый текст – случайная величина M и ключ K преобразуются в шифротекст C так, что открытый текст однозначно восстанавливается по C и K – $H(M|KC) = 0$.

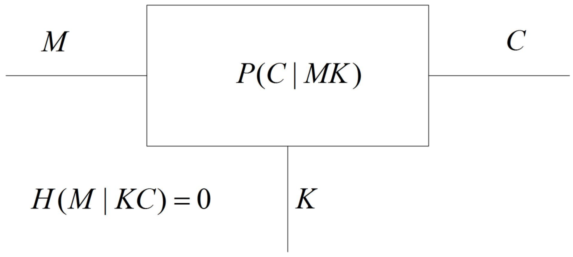

Рис. 3. Схема щифрования

Докажите границу

$$I(M,C) \geq H(M) - H(K).$$

Решение

Чтобы доказать

$$I(M,C) = H(M) - H(M|C) \geq H(M) - H(K),$$

достаточно проверить, что $H(K) \geq H(M|C)$. Рассмотрим $H(MK|C)$. С одной стороны

$$H(MK|C) = H(M|C) + H(K|MC),$$

а с другой

$$H(MK|C) = H(K|C) + H(M|KC) = H(K|C) + 0 = H(K|C).$$

Имеем

$$H(K) \geq H(K|C) = H(M|C) + H(K/MC) \geq H(M|C).$$

3. Кодирование источника с потерями 26-1

3.1. Общая граница Чебышева

Пусть $\varphi(x)$ неотрицательна ($\varphi(x) \geq 0$) и такова, что из $x > a$ вытекает $\varphi(x) > \varphi(a)$. Покажете, что для любой случайной величины $X = \{x\}$ с матожиданием $E(\varphi(X))$ справедлива оценка:

$$P(x \geq a) \leq \frac{E(\varphi(X))}{\varphi(a)}.$$

Решение

$$E(\varphi(x)) = \sum_x P(x)\varphi(x) \geq \sum_{x\geq a} P(x)\varphi(x) \geq \varphi(a)\sum_{x\geq a} P(x) =$$
$$= \varphi(a)P(x \geq a).$$

3.2. Элементарная граница Чебышева 27-1

Пусть $X = \{x\}$ – неотрицательная случайная величина с матожиданием $E(X)$. Покажите, что

$$P(x \geq a) \leq \frac{E(X)}{a}.$$

Приведите пример ситуации, когда эта граница достигается.

Решение

В общей границе достаточно выбрать $\varphi(x) = x$ при $x > 0$. Иначе,

$$P(x \geq a) = \sum_{x\geq a} P(x) \leq \sum_{x\geq a} \frac{x}{a} P(x) \leq \sum_x \frac{x}{a} P(x) = \frac{E(X)}{a}.$$

Пусть средний рост человека $E[h]$ составляет 2 метра. Граница Чебышева говорит, что вероятность встретить человека высотой ≥ 20 метров не превышает 0.1:

$$P(h \geq 20) \leq \frac{E(h)}{20} = \frac{2}{20} = \frac{1}{10}.$$

Граница достигается, когда из 10 человек 9 имеют нулевую высоту, а один – высоту в 20 метров. Тогда $E[h] = 2$ и $P(h \geq 20) = \frac{1}{10}$.

3.3. Граница Чебышева для дисперсии 28-1

Пусть $X = \{x\}$ – случайная величина с матожиданием $E(X) = m$ и дисперсией $E[(X-m)^2] = \sigma^2$. Покажите что,

$$P(|x - m| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Решение Рассмотреть случайную величину $Y = |X - m|$ и в общей границе выбрать $\varphi(y) = y^2 = |y|^2$.

3.4. Граница Чебышева для дисперсии суммы 29-1

Пусть $S_N = \frac{\sum_{k=1}^N x_k}{N}$ сумма одинаковых независимых случайных величин X с матожиданием $E[X] = m$ и дисперсией $E[(X-m)^2] = \sigma^2$. Покажите, что

$$P(|S_N - m| \geq a) \leq \frac{\sigma^2}{Na^2}.$$

Выведите отсюда следствие:

$$P\Big(|S_N - m| \geq \frac{C}{\sqrt{N}}\Big) \leq \frac{\sigma^2}{C^2}.$$

Решение Заметив, что $E[S_N] = m$, $E[(S_N - m)^2] = \frac{\sigma^2}{N}$, применить границу для дисперсии.

Пусть $X = \{x\}$ – случайная величина с матожиданием $E(X) = m$ и дисперсией $E[(X-m)^2] = \sigma^2$. Покажите что,

$$P(|x - m| \geq a) \leq \frac{\sigma^2}{a^2}.$$

3.5. Граница Чебышева для суммы двоичных величин 30-1

Пусть $X = \{0,1\}$ – двоичная случайная величина с $P(1) = p$, $P(0) = 1 - p$. Покажите, что для суммы независимых случайных величин этого рода справедлива оценка:

$$P\Big(\Big|\frac{\sum_{k=0}^{N} x_k}{N} - p\Big| \geq a\Big) = P\Big(|S_N - p| \geq a\Big) \leq \frac{p(1-p)}{Na^2}.$$

В частности,

$$P\Big(|S_N - p| \geq \frac{C}{\sqrt{N}}\Big) \leq \frac{p(1-p)}{C^2}.$$

Решение Имеем

$$m = E[X] = 1p + 0(1-p) = p.$$

$$\sigma^2 = E[(X-p)^2] = p(1-p)^2 + (1-p)p^2 = p(1-p).$$

Остальное дает граница Чебышева для суммы.

3.6. Граница Чернова 31-1

Вывести границу Чернова:

$$P(x \geq a) \leq \min_{\mu \geq 0}(e^{-\mu a} E[e^{\mu x}]).$$

Показать, что для гауссовского распределения с плотностью $\varrho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ граница Чернова дает:

$$P(x \geq a) \leq e^{-\frac{a^2}{2}}.$$

Решение

Граница Чернова – это просто общая граница Чебышева с $\varphi(x) = e^{\mu x}$, $\mu > 0$:

$$P(x \geq a) \leq \frac{E[\varphi(x)]}{\varphi(a)} = \frac{E[e^{\mu x}]}{e^{\mu a}}.$$

Для гауссовского распределения $E[e^{\mu x}]$ вычисляется:

$$E[e^{\mu x}] = \frac{1}{\sqrt{2\pi}} \int e^{-\frac{x^2}{2}} e^{\mu x} dx = e^{\frac{\mu^2}{2}}.$$

Это дает

$$P(x \geq a) \leq e^{-\mu a} e^{\frac{\mu^2}{2}}.$$

Выбор $\mu = a$ дает результат.

3.7. К конструкции множества типичных блоков 32-1

Пусть $X_N = (X_1, \ldots X_N)$ – блок независимых случайных величин с энтропией $H = H(X_k)$ каждая. Показать, что

$$P\Big(\Big|\frac{1}{N} \log \frac{1}{P(x_1 \ldots x_n)} - H\Big| \geq \beta\Big) \leq \frac{\sigma^2}{N\beta^2},$$

где

$$\sigma^2 = E[(\log \frac{1}{P(x)} - H)^2]$$

Решение

Рассмотрим случайную величину $\log \frac{1}{P(x)}$ с матожиданием H и дисперсией σ^2. К результату приводит применение границы Чебышева для суммы

$$S_N = \frac{1}{N} \sum_k \log \frac{1}{P(x_k)} = \frac{1}{N} \log \frac{1}{P(x_1 \ldots x_N).}$$

3.8. Множество типичных блоков 33-1

Пусть $X_N = (X_1, \dots X_N)$ – блок независимых случайных величин с энтропией $H = H(X_k)$ каждая. Показать, что в пространстве значений $X_N = \{\bar{x} = (x_1, \dots, x_N)\}$ можно выбрать подмножество типичных блоков

$$T_\beta(N) = \{\bar{x} : \Big|\frac{1}{N}\log\frac{1}{P(\bar{x})} - H\Big| \le \beta\}$$

с вероятностями, «почти» равными 2^{-NH} в смысле

$$2^{-N(H+\beta)} \le P(\bar{x}) \le 2^{-N(H-\beta)}$$

и при этом

$$P(T_\beta(N)) \ge 1 - \frac{1}{N\beta^2} \to 1.$$

Решение

$T_\beta(N)$ – это дополнение до множества блоков, удовлетворяющих границе Чебышева

$$P\Big(\Big|\frac{1}{N}\log\frac{1}{P(\bar{x})} - H\Big| \ge \beta\Big) \le \frac{\sigma^2}{N\beta^2},$$

Следовательно, его вероятностная мера превышает $1 - \frac{\sigma^2}{N\beta^2}$.

3.9. Оценки мощности множества типичных блоков 34-1

Пусть $X_N = (X_1, \dots X_N)$ – блок независимых случайных величин с энтропией $H = H(X_k)$ каждая, $X^N = \{\bar{x} = (x_1, \dots, x_N)\}$ – пространство их значений,

$$T_\beta(N) = \{\bar{x} : \Big|\frac{1}{N}\log\frac{1}{P(\bar{x})} - H\Big| \le \beta\}$$

множество типичных блоков в нем. Вывести границы для мощности множества $T_\beta(N)$:

$$(1 - \frac{\sigma^2}{N\beta^2})2^{N(H-\beta)} \le |T_\beta(N)| \le 2^{N(H+\beta)}$$

Решение

Вероятности типичных блоков с двух сторон «зажаты» границами

$$2^{-N(H+\beta)} = P_{min} \le P(\bar{x}) \le P_{max} = 2^{-N(H-\beta)}.$$

Ясно, что

$$|T_\beta|P_{min} \le 1; \quad |T_\beta|P_{max} \ge P(T_\beta) \ge (1 - \frac{\sigma^2}{N\beta^2});$$

3.10. Прямая теорема кодирования источника 35-1

Покажите, что во множестве $X^N = \{\bar{x} = (x_1, \dots, x_N)\}$ блоков независимых символов от источника с энтропией H можно выбрать подмножество S_δ с $P(S_\delta) \ge 1 - \delta$ мощности не превышающей $2^{N(H+\epsilon)}$, что обеспечит возможность кодирования с скоростью не более $N(H+\epsilon)$ битов на блок при вероятности ошибки не более δ.

Решение

Достаточно взять множество типичных блоков T_β с $P(T_\beta) \ge (1 - \frac{\sigma^2}{N\beta^2}) = 1 - \delta$. Получится

$$|S_\delta| \le |T_\beta| \le 2^{N(H+\beta)}$$

Примем $\beta = \epsilon$. Выбором достаточно большого N можно добиться выполнения условия $\frac{\sigma^2}{N\beta^2} < \delta$.

3.11. Обращение теоремы кодирования источника 36-1

Пусть $X^N = \{\bar{x} = (x_1, \dots, x_N)\}$ множество блоков независимых символов от источника с энтропией . Покажите, что при любом выборе подмножества S_δ мощности $|S_\delta| \leq 2^{N(H-\epsilon)}$ вероятностная мера этого множества $P(S_\delta)$ стремится к нулю при $N \to \infty$.

Решение

Выберем некоторое множество типичных блоков T_β и пусть $\bar{T}_\beta$ – его дополнение. Представив S_δ в виде объединения непересекающихся множеств

$$S_\delta = (S_\delta \cap T_\beta) \cup (S_\delta \cap \bar{T}_\beta),$$

найдем

$$P(S_\delta) = P(S_\delta \cap T_\beta) + P(S_\delta \cap \bar{T}_\beta) \leq |S_\delta| 2^{-N(H-\beta)} + P(\bar{T}_\beta) \leq$$

$$\leq 2^{N(H-\epsilon)} 2^{-N(H-\beta)} + \frac{\sigma^2}{N\beta^2} = 2^{-N(\epsilon-\beta)} + \frac{\sigma^2}{N\beta^2}$$

Выбрав $\beta < \epsilon$, получим стремление к нулю $P(S_\delta)$ при больших N.

4. Кодирование источника без потерь 37-1

4.1. Неравенство Крафта

Пусть l_j – набор длин слов двоичного кода с однозначным декодированием. Докажите неравенство Крафта:

$$S = \sum_j 2^{-l_j} \leq 1.$$

Решение

$$S^N = \left(\sum_j 2^{-l_j}\right)^N = \sum_{j_1, \dots, j_N} 2^{-(l_{j_1} + \dots + l_{j_N})} = \sum_{l=Nl_{min}}^{Nl_{max}} A(l) 2^{-l},$$

где l_{min}, l_{vax} – минимальная и максимальная длина слова, а $A(l)$ – число вариантов разбиения всех возможных двоичных l-блоков на N кодовых слов всех. Всего имеется 2^l двоичных l-блоков. Если код однозначно декодируемый, то каждый из них допускает не более одного разбиения но слова. Поэтому $A(l) \leq 2^l$ и

$$S^N \leq \sum_{l=Nl_{min}}^{Nl_{max}} 1 \leq Nl_{max}.$$

Так что S^N не уходит в бесконечность с ростом N. Но это возможно только при $S \leq 1$.

4.2. Неравенство Крафта 38-1

Пусть l_j – набор длин слов q-ичного кода с однозначным декодированием. Докажите неравенство Крафта:

$$S = \sum_j q^{-l_j} \leq 1.$$

4.3. Нижняя граница для длины двоичного префиксного кода 39-1 1-2

Пусть M символов источника с вероятностями p_m закодированы двоичным префиксным кодом с длинами слов l_m. Докажите нижнюю границу для средней длины кодового слова:

$$L = \sum_m p_m l_m \geq H,$$

где $H = -\sum_m p_m \log p_m$ – энтропия источника.

Решение

Пусть

$$\gamma = \sum_m 2^{-l_m}.$$

Согласно неравенству Крафта, $\gamma \leq 1$. Введем распределение вероятностей $q_m = \frac{2^{-l_m}}{\gamma}$. Имеем

$$L-H = \sum_m p_m \log 2^{l_m} p_m = \sum_m p_m \log \frac{p_m}{2^{-l_m}} = \sum_m p_m \log \frac{p_m}{\gamma q_m} = D(P||Q)+\log\frac{1}{\gamma} \geq \log\frac{1}{\gamma} \geq 0.$$

4.4. Нижняя граница для длины q-ичного префиксного кода 40-1 2-2

Пусть M символов источника с вероятностями p_m закодированы q-ичным префиксным кодом с длинами слов l_m. Докажите нижнюю границу для средней длины кодового слова:

$$L = \sum_m p_m l_m \geq \frac{H}{\log q},$$

где $H = -\sum_m p_m \log p_m$ – энтропия источника.

Решение

Пусть $H_q = -\sum_m p_m \log_q p_m$ – q-ичная энтропия источника. Из предыдущего ясно, что

$$L \geq H_q = \frac{H}{\log q},$$

поскольку $\log_q a = \log_q 2^{\log a} = \log a \log_q 2 = \frac{\log a}{\log q}$.

4.5. Верхняя граница для длины двоичного префиксного кода 3-2

Покажите, что для M-символьного источника с распределением вероятностей p_m и энтропией H существует двоичный префиксный код со средней длиной слова, удовлетворяющей границе:

$$L \leq \sum_m p_m l_m \leq H + 1.$$

Решение

Выберем $l_m = \log\frac{1}{p_m} + \mu_m$, где $0 \leq \mu_m < 1$ дополнение $\log\frac{1}{p_m}$ до ближайшего целого сверху. Набор длин l_m удовлетворяет неравенству Крафта:

$$\sum_m 2^{-l_m} = \sum_m p_m 2^{-\mu_m} \leq \sum_m p_m = 1.$$

Поэтому существует префиксный код с эти набором длин и для него

$$L = \sum_m p_m l_m = \sum_m p_m \log\frac{1}{p_m} + \sum_m p_m \mu_m = H + \sum_m p_m \mu_m \leq 1$$

4.6. Верхняя граница для длины q-ичного префиксного кода 4-2

Покажите, что для M-символьного источника с распределением вероятностей p_m и энтропией H существует q-ичный префиксный код со средней длиной слова, удовлетворяющей границе:

$$L \le \sum_m p_m l_m \le \frac{H}{\log q} + 1.$$

4.7. Уточненная верхняя граница для длины двоичного префиксного кода 5-2

Покажите, что для M-символьного источника с распределением вероятностей p_m и энтропией H существует двоичный префиксный код со средней длиной слова, удовлетворяющей границе:

$$L \le \sum_m p_m l_m \le H + \mu,$$

где $\mu = \sum_m p_m \mu_m$ среднее значение дополнения логарифма $\log \frac{1}{p_m}$ до ближайшего целого сверху: $l_m = \log \frac{1}{p_m} + \mu_m$, $0 \le \mu_m < 1$.

Решение

См. доказательство верхней границы.

4.8. Случай равенства средней длины и энтропии 6-2

Покажите, что если вероятности p_m всех M символов источника выражаются степенями двойки ($p_m = 2^{-l_m}$), то существует двоичный префиксный код со средней длиной, равной энтропии источника H: $L = \sum_m p_m l_m = H$. Покажите также, что набор l_m длин этого кода удовлетворяет равенству Крафта:

$$\sum_m 2^{-l_m} = 1.$$

Решение

В представлении $l_m = \log \frac{1}{p_m} + \mu_m$ для длин слов кода все поправки μ_m равны 0.

4.9. Конструкция двоичного кода Хаффмана 7-2

Постройте оптимальный двоичный код Хаффмана для источника (a, b, c, d, e) с вероятностями символов $(0.25, 0.25, 0.2, 0.15, 0.15)$.

Решение

Объединяем d и e в виртуальный символ (d, e) с вероятностью $0.15{+}0.15 = 0.3$. Далее объединяем, скажем, b и c в (b, c) с вероятностью $0.25{+}0.2 = 0.45$. Далее, объединяем a и (d, e) в (a, d, e) с вероятностью $0.25{+}0.3 = 0.55$. Наконец, объединяем (a, d, e) и bc.

4.10. Конструкция двоичного кода Хаффмана 8-2

Постройте оптимальный код Хаффмана для кодирования восьми двоичных 3-блоков из независимых символов с вероятностью единицы $p = 1/3$. Оцените среднюю длину слова.

Решение

Имеем символ (111) с вероятностью $1/27$, три символа $(110), (101), (011)$ с вероятностями $2/27$, три символа $(100), (010), (001)$ с вероятностями $4/27$ и символ (000) вероятностью $8/27$. Конструкция кода показана на рисунке.

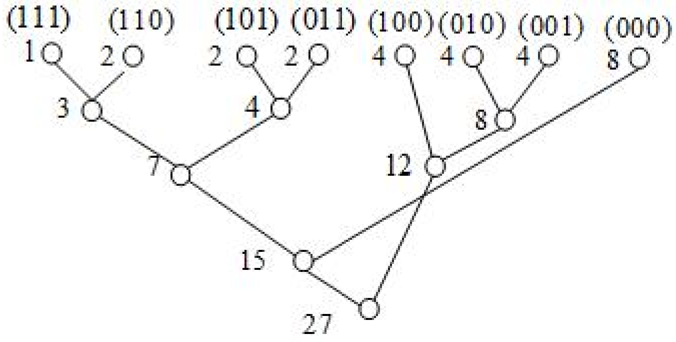

Рис. 4. Дерево Хаффмана

4.11. Конструкция двоичного кода Хаффмана 9-2

Постройте оптимальный код Хаффмана для кодирования восьми двоичных 3-блоков из независимых символов с вероятностью единицы $p = 1/4$. Оцените среднюю длину слова.

Решение

Имеем символ (111) с вероятностью $1/64$, три символа $(110), (101), (011)$ с вероятностями $3/64$, три символа $(100), (010), (001)$ с вероятностями $9/64$ и символ (000) вероятностью $27/64$. Конструкция кода показана на рисунке.

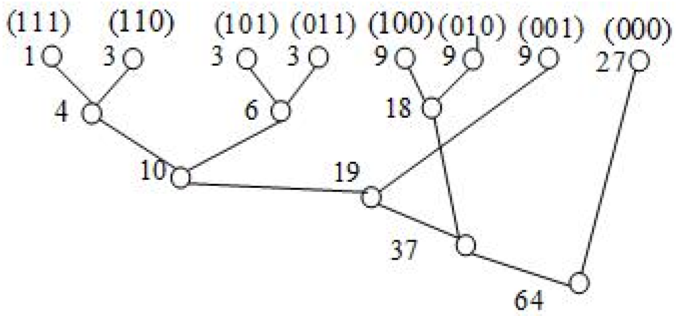

Рис. 5. Дерево Хаффмана

4.12. К двоичному коду Хаффмана

Обратившись к конструкции кода Хаффмана для кодирования восьми двоичных 3-блоков из независимых символов с вероятностью единицы p, покажите, что ни при каком значении p длина оптимального кода Хаффмана не может достигать 7.

Решение В коде для 8-и символов существует слово длины 7, если два листа имеется только не последнем, седьмом ярусе. Пусть, для конкретности, $p < 1/2$. Тогда вероятности символов упорядочены как

$$p^3 < p^2(1-p) < p(1-p)^2 < p^3.$$

На верхнем ярусе происходит слияние символов с вероятностями p^3 и $p^2(1-p)$. Это дает виртуальный символ с вероятностью $p^3 + p^2(1-p) = p^2$. Но $p^2 < p^2(1-p)$. Поэтому на следующем этапе с необходимостью сливаются два из трех имеющихся листьев с вероятностью $p^2(1-p)$. А это исключает присутствие слова максимально возможной длины 7.

4.13. К проблеме единственности кода Хаффмана 10-2

Для 4-символьного источника (a, b, c, d) с вероятностями $(\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{3})$ постройте все различные оптимальные кода Хаффмана. Проверьте факт совпадения их длин.

Решение

После объединения символов (a, b) в один комбинированный остается три символа $((a, b), c, d)$ с равными вероятностями. Имеется три варианта выбора следующей пары.

4.14. Полные коды Хаффмана 11-2

Префиксный код назовем полным, если в его дереве отсутствуют свободные листья, то есть, неравенство Крафта выполняется с равенством. При каких размерах M алфавита источника существуют полные q-ичные префиксные коды ?

Решение

Число листьев полного q-ичного дерева выражается формулой: $M = q + s(q-1)$ – имеется q узлов на первом – выходящем из корня уровне. Расщепление каждого из узлов исключает один и добавляет q новых. Всего при каждом расщеплении добавляется $(q-1)$ узел.

4.15. Полные коды Хаффмана 12-2

Префиксный код назовем полным, если в его дереве отсутствуют свободные листья, то есть, неравенство Крафта выполняется с равенством. Покажите, что полный двоичный префиксный код существует при любом размере $M \geq 2$ алфавита источника.

Решение

При $q = 2$ число листьев полного q-ичного дерева выражается формулой: $M = 2 + s(2-1) = 2 + s$ и может быть любым с $M \geq 2$.

4.16. Элементарный код Танстолла 13-2

Пусть двоичный префиксный код Танстолла со словами $(0, 10, 11)$ используется для преобразования потока равновероятных двоичных символов источника в троичный алфавит (A, B, C). Найти энтропию на символ тричного выходного потока и сравнить ее со средним числом битов источника на троичный символ.

Решение Символы (A, B, C), очевидно, статистически независимы и встречаются с вероятностями $(1/2, 1/4, 1/4)$. Энтропия этого распределения составляет $H = \frac{1}{2}\log 2 + 2\frac{1}{4}\log 4 = \frac{3}{2}$ бита на символ. Это совпадает со средним числом битов на троичный символ.

4.17. Код Танстолла 14-2

Построить двоичный префиксный код Танстолла для кодирования потока незавичимых двоичных символов с вероятностью единицы $p = 1/3$ на девять 2-блоков символов троичного алфавита (A, B, C) Оценить среднее число битов на троичный 2-блок.

Решение Дерево строим снизу вверх, расщепляя на каждом шаге наиболее вероятный лист на левый и правый с условными вероятностями $(1/3, 2/3)$.

В результате получается дерево со следующим набором вероятностей листьев (слева направо):

$$\frac{27, 18, 36, 18, 36, 36, 24, 16, 32}{243}$$

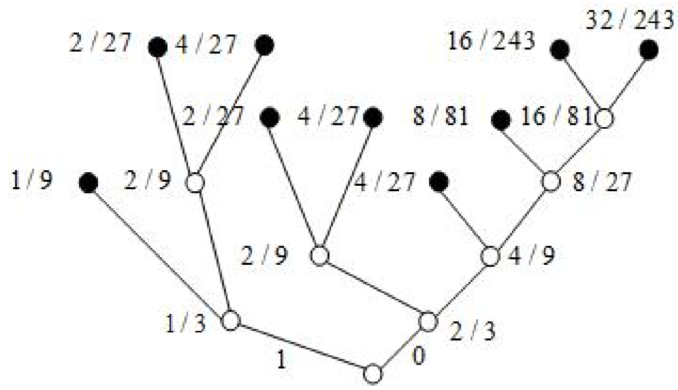

Рис. 6. Дерево Танстолла

4.18. Арифметическое кодирование

Пусть арифметический кодер преобразует поток независимых двоичных символов с вероятностью нуля, равной $1/3$ в двоичный же выходной поток. На вход кодера уже поступили два символа – $(0,0),(0,1),(1,0)$ или $(1,1)$. Для каждого из этих четырех вариантов укажите: Какие символы уже могут быть посланы на выход ? Какие символы следует добавить, чтобы оборвать процесс кодирования с возможностью восстановления переданной пары ?

Решение

Парам закодированных символов отвечают показанные на рисунке интервалы с границами в точках $0, 1/9, 3/9, 5/9, 1$.

1. Интервал (00) длины $1/9$ вложен в цилиндр [000]. Три эти бита уже можно послать на выход. Чтобы обеспечить однозначное декодирование нужно сообщить получателю номер цилиндра, лежащего внутри этого интервала. Таковым является цилиндр [0000] длины $1/16$. После трех уже посланных на выход нулей достаточно добавить еще один.

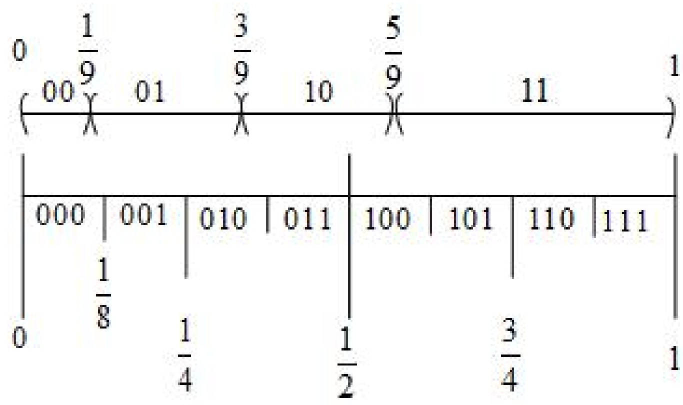

Рис. 7. Дерево Танстолла

2. Интервал (01) длины $2/9$ вложен в цилиндр [0] длины $1/2$. На выход можно послать только это бит. Внутри интревала лежит цилиндр [001]. Достаточно дополнить уже переданный нуль парой дополнительных битов 01.

3. Интервал (10) длины 2/9 не вложен не в какой из цилиндров. На выход нельзя послать ничего. Для однозначной идентификации следует послать три избыточные бита интервала [011].

4. Интервал (11) длины 4/9 вложен в цилиндр [1] длины 1/2. Бит 1 можно послать на выход. В интервал вложен цилиндр [11] длины 1/4 – достаточно послать на выход еще одну избыточную единицу.

4.19. Арифметическое кодирование 2

Пусть на вход арифметического кодера, преобразующего поток независимых равновероятных троичных символов A, B, C в двоичный выходной поток поступили два символа. Для каждой из девяти возможных пар укажите двоичный код, посланный на выход кодера.

Решение

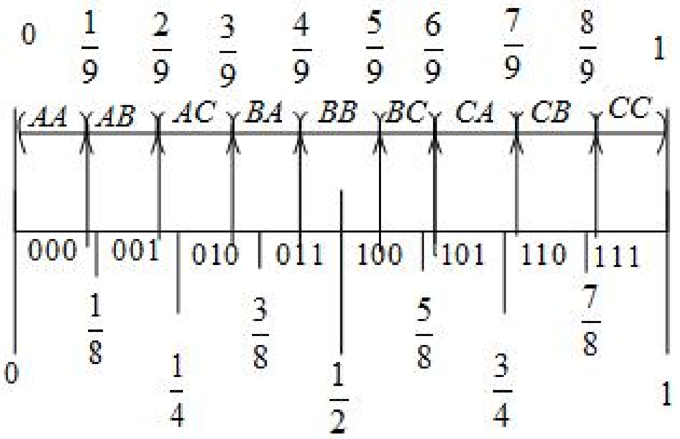

Рис. 8. Дерево Танстолла

Расположение границ девяти интервалов длины 1/9 относительно границ двоичных цилиндров с длинами 1/2, 1/4 и 1/8 показано на рисунке. На выход кодера можно сообщить идентификационный код цилиндра, целиком содержащего интервал. Таковыми являются: для (AA) - код 000, для (AB) - код 00, для (AC) - код 0, для (BA) - код 01, для (BB) – пустой код, для (BC) - код 10, для (CA) - код 1, для (CB) - код 11, для (CC) - код 111.

5. Байесовское оценивание

5.1. Байесовская оценка вероятности 17-2

Пусть в последовательности из $n = 100$ бросаний несимметричной монеты с неизвестной вероятностью p выпадения орла выпало $k = 50$ орлов и $n - k = 50$ решек. По результату наблюдения k построить байесовскую оценку плотности $\rho(p)$ апостериорного распределения параметра p, считая априорное распределение $\rho_0(p)$ равномерным. Найти матожидание $E[p]$ параметра p по плотности $\rho(p)$ – эмпирическую оценку вероятности p. Принять во внимание, что

$$\int_0^1 p^n(1-p)^m dp = \frac{n!m!}{(n+m+1)!}.$$

Решение

По формуле Байеса
$$\varrho(p|k) = \frac{P(k|p)\varrho_0(p)}{P(k)},$$
где $P(k|p) = \left(\begin{array}{c} n \\ k \end{array}\right) p^k(1-p)^{n-k}$, а
$$P(k) = \int_0^1 P(k|p)\rho_0(p)dp = \left(\begin{array}{c} n \\ k \end{array}\right) \int_0^1 p^k(1-p)^{n-k}dp = \left(\begin{array}{c} n \\ k \end{array}\right) \frac{k!(n-k)!}{(n+1)!}$$
Так что
$$\varrho(p|k) = \frac{(n+1)!}{k!(n-k)!} p^k(1-p)^{n-k}.$$

Далее,
$$E[p] = \int_0^1 p\varrho(p|k)dp = \frac{(n+1)!}{k!(n-k)!}\frac{(k+1)!(n-k)!}{(n+2)!} = \frac{k+1}{n+2}.$$
При $n = 100$, $k = 50$ $E[p] = 1/2$.

5.2. Лемма о совместительстве 15-2

Пусть на Ваш выбор предложено N мест работы с окладами S_n, $n = [1, N]$. Допускается частичная занятость с выплатой $q_n S_n$, пропорциональной доле q_n рабочего времени. Требуется распределить рабочее время (выбрать набор долей q_n с $\sum_n q_n = 1$) так, чтобы максимизировать суммарный доход
$$S = \sum_{n=1}^{N} q_n S_n.$$

Решение

Очевидно, что следует выбрать место работы с максимальным окладом и проводить там все рабочее время:
$$q_n = 1 \quad \text{при} \quad n = \mathrm{argmax_n(S_n)}.$$

5.3. Оценка максимума апостериорной вероятности 16-2

Пусть решения относительно значений символа $X = \{x\}$ выносятся по результату наблюдения значения $Y = \{y\}$ на выходе канала с матрицей условных вероятностей $P_{Y|X}(y|x) = P(y|x)$. Покажите, что решение по максимуму апостериорной вероятности
$$\tilde{x} = \mathrm{argmax_x P_{X|Y}(x|y)}$$
минимизирует среднюю вероятность ошибки
$$P_e = \sum_{\tilde{x}\neq x} P(x, \tilde{x}).$$

Решение

Будем принимать решения случайно согласно набору условных распределений $Q(\tilde{x}|y)$. Имеем
$$1 - P_e = \sum_{\tilde{x}=x} P(x, \tilde{x}) = \sum_{x,y} P(x)P(y|x)Q(x|y) = \sum_{y} P(y) \sum_{x} P(x|y)Q(x|y)$$

Согласно лемме о совместительстве, каждая из внутренних сумм по x максимизируется распределением $Q(x|y)$, принимающим значение 1 при

$$\tilde{x} = \mathrm{argmax_x P(x|y)}.$$

5.4. Оценка максимума правдоподобия 18-2

Пусть решения относительно значений символа $X = \{x\}$ выносятся по результату наблюдения значения $Y = \{y\}$ на выходе канала с матрицей условных вероятностей $P_{Y|X}(y|x) = P(y|x)$. Покажите, что при равномерном априорном распределении ($P(x) = \frac{1}{|X|}$) оценка по максимуму апостериорной вероятности

$$\tilde{x} = \mathrm{argmax_x P_{X|Y}(x|y)}$$

сводится к оценке по максимуму правдоподобия:

$$\tilde{x} = \mathrm{argmax_x P_{Y|X}(y|x)}.$$

Решение
Имеем:

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}.$$

При $P(x) = const$ максимизация по x апостериорной вероятности $P(x|y)$ эквивалента максимизации функции правдоподобия $P(y|x)$.

5.5. Сложение отношений правдоподобия 19-2

Пусть решения относительно значений двоичного символа $X = \{0, 1\}$ с вероятностью $P(x = 1) = q$ выносятся по результату наблюдения значения $Y = \{y\}$ на выходе канала с матрицей условных вероятностей $P_{Y|X}(y|x) = P(y|x)$. Покажите, что логарифм отношения правдоподобия для апостериорного распределения представляется суммой

$$\ln\frac{P(x=0|y)}{P(x=1|y)} = \ln\frac{P(y|x=0)}{P(y|x=1)} + \ln\frac{1-q}{q}.$$

Решение Имеем

$$P(x=0|y) = \frac{P(y|x=0)P(x=0)}{P(y)}; \quad P(x=1|y) = \frac{P(y|x=1)P(x=1)}{P(y)}$$

5.6. Сложение отношений правдоподобия независимых наблюдений 20-2

Пусть решения относительно значений двоичного символа $X = \{0, 1\}$ с вероятностью $P(x = 1) = q$ выносятся по результатам двух независимых наблюдений значения $Y = \{y\}$ и $Z = \{z\}$ на выходах каналов с матрицами условных вероятностей $P_{Y|X}(y|x) = P(y|x)$ и $P_{Z|X}(z|x) = P(z|x)$. Покажите, что логарифм отношения правдоподобия для апостериорного распределения представляется суммой

$$\ln\frac{P(x=0|yz)}{P(x=1|yz)} = \ln\frac{P(y|x=0)}{P(y|x=1)} + \ln\frac{P(z|x=0)}{P(z|x=1)} + \ln\frac{1-q}{q}$$

Решение Имеем

$$P(x=0|yz) = \frac{P(yz|x=0)P(x=0)}{P(yz)} = \frac{P(y|x=0)P(z|x=0)P(x=0)}{P(yz)};$$

$$P(x=1|yz) = \frac{P(yz|x=1)P(x=1)}{P(yz)} = \frac{P(y|x=1)P(z|x=1)P(x=1)}{P(yz)}.$$

5.7. Отношение правдоподобия суммы 21-2

Пусть $z = x + y$ сумма по модулю два двух случайных битов с вероятностями единиц $P(x = 1) = p$, $P(y = 1) = q$ (логарифмами отношений правдоподобия $\lambda_x = \ln \frac{1-p}{p}$ и $\lambda_y = \ln \frac{1-q}{q}$). Покажите, что логарифм отношения правдоподобия λ_z для суммы z представляется в виде:

$$\lambda_z = \ln \frac{1 + \operatorname{th}(\frac{\lambda_x}{2}) \operatorname{th}(\frac{\lambda_y}{2})}{1 - \operatorname{th}(\frac{\lambda_x}{2}) \operatorname{th}(\frac{\lambda_y}{2})}.$$

Решение

Имеем: $P = P(z = 1) = p(1 - q) + q(1 - p) = p + q - 2pq$, так что

$$1 - 2P = (1 - 2p)(1 - 2q).$$

Но $p = \frac{e^{-\lambda_x}}{1+e^{-\lambda_x}}$, а $1 - p = \frac{e^{\lambda_x}}{1+e^{\lambda_x}}$, так что

$$1 - 2p = (1 - p) - p = \operatorname{th} \frac{\lambda_x}{2}.$$

Аналогично, $1 - 2q = \operatorname{th} \frac{\lambda_y}{2}$, $1 - 2P = \operatorname{th} \frac{\lambda_z}{2}$. Так что $\operatorname{th} \frac{\lambda_z}{2} = \operatorname{th} \frac{\lambda_x}{2} \operatorname{th} \frac{\lambda_y}{2}$. Осталось воспользоваться легко проверяемым тождеством

$$2 \operatorname{th}^{-1}(u) = \ln \frac{1 + u}{1 - u}.$$

5.8. Максимум апостериорной вероятности в двоичном случае 22-2

Пусть двоичная случайная величина X с вероятностью $P(x = 1) = q$ оценивается по двоичным значениям Y на выходе двоичного симметричного канала с вероятностью ошибки $p < 1/2$. Покажите, что оценивание x по максимуму апостериорной вероятности сводится к решению $\tilde{x} = y$ при $q > p$ и решению $\tilde{x} = 0$ при $q < p$. Каковы средние вероятности ошибок в одном и другом случае.

Решение

Выпишем отношения правдоподобия для апостериорного распределения x при заданном y:

$$\frac{P(x = 1|y = 0)}{P(x = 0|y = 0)} = \frac{P(y = 0|x = 1)p(x = 1)}{P(y = 0|x = 0)p(x = 0)} = \frac{p}{(1 - p)} \frac{q}{(1 - q)};$$

$$\frac{P(x = 1|y = 1)}{P(x = 0|y = 1)} = \frac{P(y = 1|x = 1)p(x = 1)}{P(y = 1|x = 0)p(x = 0)} = \frac{(1 - p)}{p} \frac{q}{(1 - q)}.$$

По наблюдении $y = 0$ следует выносить решение $\tilde{x} = 0$, если $(1 - p)(1 - q) > pq$ или $q < 1 - p$ и решение $\tilde{x} = 1$ при $q > 1 - p$.

По наблюдении $y = 1$ следует выносить решение $\tilde{x} = 1$, если $(1 - p)q > q(1 - p)$ или $q > p$ и решение $\tilde{x} = 0$ при $q < p$.

При $q < p$ декодер по максимуму апостериорной вероятности выносит решение $\tilde{x} = 0$ независимо от y. При $p < q < 1 - p$ решение совпадает с принятым символом $\tilde{x} = y$. При $q > 1 - p$ неизменно выносится решение $\tilde{x} = 1$.

Выпишем совместное распределение $P(x, y)$:

$$p(0, 0) = (1 - q)(1 - p); \quad P(0, 1) = (1 - q)p; \quad P(1, 0) = qp; \quad P(1, 1) = q(1 - p).$$

Видно, что при вынесении решения по правилу $\tilde{x} = y$ средняя вероятность ошибки составляет

$$Pe = P(1,0) + P(0,1) = p$$

При вынесении решения по правилу $\tilde{x} = 0$ средняя вероятность ошибки составляет

$$Pe = P(1,1) + P(1,0) = q,$$

что лучше p при $q < p$. При вынесении решения по правилу $\tilde{x} = 1$ средняя вероятность ошибки составляет

$$Pe = P(0,1) + P(0,0) = 1 - q,$$

что лучше p при $q > 1 - p$. Таким образом, утруждать себя наблюдением выходов канала имеет смысл только при $p < q < 1 - p$.

5.9. Отношение правдоподобия и двоичная случайная величина в гауссовском шуме 23-2

Пусть двоичная случайная величина $X = 0, 1$ наблюдается на выходе гауссовского канала с условными плотностями

$$\varrho(y|0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-c)^2}{2\sigma^2}}; \quad \varrho(y|1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+c)^2}{2\sigma^2}};$$

Пусть

$$\lambda(y) = \ln \frac{\varrho(y|0)}{\varrho(y|1)}$$

– логарифм отношения правдоподобия. Покажите, что

$$\lambda(y) = \frac{2yc}{\sigma^2}, \quad \varrho(y|0) = \frac{1}{1 + e^{-\lambda}}; \quad \varrho(y|1) = \frac{1}{1 + e^{\lambda}};$$

Решение Простые преобразования.

5.10. Оценивание двоичной случайной величины в гауссовском шуме 24-2

Пусть двоичная случайная величина $X = 0, 1$ с $P(x = 1) = q$ наблюдается на выходе гауссовского канала с условными плотностями

$$\varrho(y|0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-c)^2}{2\sigma^2}}; \quad \varrho(y|1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+c)^2}{2\sigma^2}};$$

Предложите правила оценивания X по наблюдению y по максимуму апостериорной вероятности и максимуму правдоподобия.

Решение Для логарифма апостериорного отношения правдоподобия имеем:

$$\ln \frac{P(x = 0|y)}{P(x = 1|y)} = \ln \frac{\rho(y|x = 0)P(x = 0)}{\rho(y|x = 1)P(x = 1)} = \frac{2cy}{\sigma^2} + \ln \frac{1 - q}{q}.$$

Решение максимального правдоподобия (случай $q = 1/2$) выносится по знаку y: $x = 0$ если $y > 0$. Решение максимального правдоподобия – по результату сравнения $\frac{2cy}{\sigma^2}$ с порогом $\ln \frac{1-q}{q}$.

6. Пропускная способность, теоремы кодирования

6.1. Граница Фано26-2

Пусть значения K-значной случайной величины X оценивается по наблюдениям Y на выходе канала с матрицей условных вероятностей $P(Y|X)$. Покажите, что при любом алгоритме оценивания $Q(\tilde{X}|Y)$ вероятность ошибки

$$P_e = \sum_{x \neq \tilde{x}, y} P(x)P(y|x)Q(\tilde{x}|y)$$

удовлетворяет границе Фано:

$$h(P_e) + P_e \log_2(K-1) \geq H(X|\tilde{X}),$$

где $h(x) = -x \log x - (1-x)\log(1-x)$ - двоичная энтропия.

Решение:

Введем случайную величину $U(x, \tilde{x})$, равную 1 при $x = \tilde{x}$ и нулю в противном случае и рассмотрим $H(XU|\tilde{X})$. С одной стороны,

$$H(XU|\tilde{X}) = H(X|\tilde{X}) + H(U|X, \tilde{X}) = H(X|\tilde{X}),$$

поскольку условия $X, \tilde{X}$ вполне определяют значение U. А с другой

$$H(XU|\tilde{X}) = H(U|\tilde{X}) + H(X|U\tilde{X}) \leq H(U) + P(u=0)H(X|u=0, \tilde{X}) + P(u=1)H(X|u=1, \tilde{X}) =$$

.

$$\leq H(U) + P_e H(X|x \neq \hat{x}) = h(P_e) + P_e \log(K-1).$$

Более доказывать нечего.

6.2. Двоичная граница Фано 27-2

Пусть значения двоичной случайной величины $X = 0, 1$ оценивается по наблюдениям Y на выходе канала с матрицей условных вероятностей $P(Y|X)$. Покажите, что при любом алгоритме оценивания $Q(\tilde{X}, Y)$ вероятность ошибки

$$P_e = \sum_{x \neq \tilde{x}, y} P(x)P(y|x)Q(\tilde{x}|y)$$

удовлетворяет границе Фано:

$$h(P_e) \geq H(X|\tilde{X}),$$

где $h(x) = -x \log x - (1-x)\log(1-x)$ - двоичная энтропия.

Решение: Частный случай общей границы при $K = 2$.

6.3. К марковским цепям 1 25-2

Пусть случайные величины $X \to Y \to Z$ образуют цепь маркова в том смысле, что

$$P(X, Y, Z) = P(X)P(Y|X)P(Z|Y).$$

Покажите что $H(Z|X, Y) = H(Z|Y)$, $H(X|Y, Z) = H(X|Y)$.

Решение: Первое – это следствие того, что для цепи Маркова $P(Z|X, Y) = P(Z|Y)$. Чтобы доказать второе, покажем, что $Z \to Y \to X$ также образуют (обращенную) цепь маркова. По Байесу

$$P(X)P(Y|X)P(Z|Y) = P(X, Y)P(Z|Y) = P(Y)P(X|Y)P(Z|Y) =$$

$$= P(X|Y)P(Y)P(Z|Y) = P(X|Y)P(Y, Z) = P(X|Y)P(Y|Z)P(Z).$$

6.4. К марковским цепям 2 28-2

Пусть случайные величины $X \to Y \to Z$ образуют цепь маркова в том смысле, что

$$P(X, Y, Z) = P(X)P(Y|X)P(Z|Y).$$

Покажите что $H(X, Z|Y) = H(X|Y) + H(Z|Y)$.

Решение: Применив формулу Байеса

$$P(X)P(Y|X)P(Z|Y) = P(X, Y)P(Z|Y) = P(Y)P(X|Y)P(Z|Y),$$

найдем, что случайные величины X, Z условно независимы при каждом данном Y. Это и дает результат.

6.5. Лемма об обработке информации 1 29-2

Пусть случайные величины $X \to Y \to Z$ образуют цепь маркова в том смысле, что

$$P(X, Y, Z) = P(X)P(Y|X)P(Z|Y).$$

Покажите что

$$I(X, Y) \geq I(X, Z).$$

Решение: Имеем

$$I(X|Y) = H(X) - H(X|Y) = H(X) - H(X|Y, Z),$$

поскольку для марковской цепи $H(X|Y) = H(X|Y, Z)$. Но $H(X|Y, Z) \leq H(X|Z)$ – введение условия снижает энтропию. Поэтому

$$I(X|Y) = H(X) - H(X|Y, Z) \geq H(X) - H(X|Z) = I(X, Z).$$

6.6. Лемма об обработке информации 2 30-2

Пусть случайные величины $X \to Y \to Z$ образуют цепь маркова в том смысле, что

$$P(X, Y, Z) = P(X)P(Y|X)P(Z|Y).$$

Покажите что

$$I(Y, Z) \geq I(X, Z).$$

Решение:

Имеем

$$I(Y, Z) = H(Z) - H(Z|Y) = H(Z) - H(Z|X, Y),$$

поскольку для марковской цепи $H(Z|Y) = H(Z|X, Z)$. Но $H(Z|X, Y) \leq H(Z|X)$ – введение условия снижает энтропию. Поэтому

$$I(Y, Z) = H(Z) - H(Z|X, Y) \geq H(Z) - H(Z|X) = I(X, Z).$$

6.7. Лемма об обработке информации 3 31-2

Пусть случайные величины $U \to X \to Y \to V$ образуют цепь маркова в том смысле, что

$$P(U, X, Y, V) = P(U)P(X|U)P(Y|X)P(V|Y).$$

Покажите что

$$I(U, V) \leq I(X, Y).$$

Решение: Тройка $U \to X \to V$ (с опущенным Y) образует цепь маркова с

$$P(V|X) = \sum_y P(Y|X)P(V|Y)$$

. Поэтому $I(U,V) \leq I(X,V)$. Но тройка $X \to Y \to V$ – это также цепь маркова с $I(X,V) \leq I(X,Y)$.

6.8. Обращение теоремы кодирования 37-2

Пусть информационные двоичные K-блоки U преобразуются кодером в слова X длины L над некоторым алфавитом, передаются по каналу с матрицей условных вероятностей $P(Y|X)$ и пропускной способностью C битов на символ. Получающиеся на выходе L-блоки Y преобразуются декодером в выходные двоичные K-блоки V, см. рисунок.

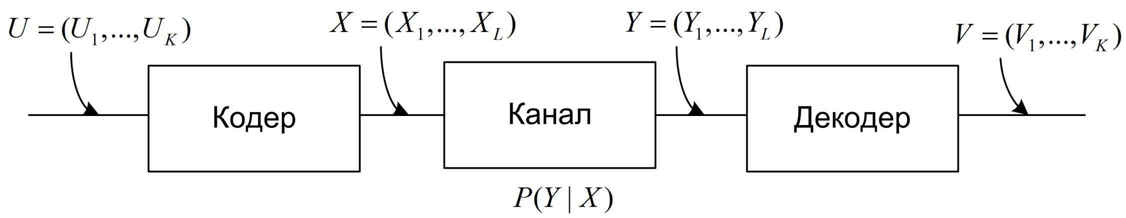

Рис. 9. Модель канала

Пусть $P_b(k) = P(U_k \neq V_k)$ – вероятность ошибки в k-ом бите, $P_b = \frac{1}{K}\sum_{k=1}^{K} P_b(k)$ – средняя вероятность ошибки на бит. Покажите, что

$$h(P_b) \geq 1 - \frac{C}{R},$$

где $R = \frac{K}{L}$ битов на использование канала, а $h(x)$ – двоичная энтропия. Постройте набросок графика этой границы.

Решение:

Опираясь на определение пропускной способности и лемму об обработке, строим нижнюю и верхнюю оценки для взаимной информации $I(X,Y)$:

$$LC \geq I(X,Y) \geq I(U,V) = H(U) - H(U|V) = K - H(U|V)$$

То есть

$$1 - \frac{C}{R} \leq \frac{1}{K}H(U|V).$$

Но

$$H(U|V) = H(U_1,...,U_K|V_1,...,V_K) = H(U_1|V_1,...,V_K) + H(U_2,...,U_K|U_1,V_1,...,V_K) \leq$$

$$\leq H(U_1|V_1) + H(U_2,...,U_K|V_2,...,V_K)$$

Продолжая действовать по этой схеме, и применив границу Фано придем к

$$H(U|V) \leq \sum_k H(U_k|V_k) \leq \sum_k h(P_b(k)$$

Таким образом,

$$1 - \frac{C}{R} \le \frac{1}{K} H(U|V) \le \frac{1}{K} \sum_k h(P_b(k)).$$

Осталось применить неравенство Йенсена для выпуклой вверх функции $h(x)$:

$$E[h(x)] \le h(E[x]) \quad \Rightarrow \quad \frac{1}{K} \sum_k h(P_b(k)) \le h\Big(\frac{1}{K} \sum_k P_b(k)\Big) = h(P_b).$$

Окончательно,

$$h(P_b) \ge 1 - \frac{C}{R}.$$

6.9. Совместно-типичные блоки 32-2

Пусть совместное распределение вероятностей $P(X, Y)$ с энтропией $H(X, Y)$ определяет совместное распределение вероятностей N-блоков (X^N, Y^N)

$$X^N = \{x^N = (x_1, x_2, \ldots, x_N);\, x_j \in X\} \quad Y^N = \{y^N = (y_1, y_2, \ldots, y_N);\, y_j \in Y\}$$

по правилу

$$P(x^N, y^N) = \prod_j P(x_j, y_j).$$

Введем множество β типичных пар

$$J_\beta = \Big\{(x^N, y^N) : \Big|\frac{1}{N} \log \frac{1}{P(x^N, y^N)} - H(X, Y)\Big| \le \beta\Big\}$$

Показать, что для мощности множества J_β справедлива оценка

$$|J_\beta| \le 2^{N(H(X,Y)+\beta)}$$

Решение: Согласно границе Чебышева

$$P\Big(\Big|\frac{1}{N} \log \frac{1}{P(x^N, y^N)} - H(X, Y)\Big| \ge \beta\Big) \le \frac{\sigma^2}{N\beta^2},$$

то есть вероятности типичных пар заключены в интервале

$$2^{-N(H(X,Y)+\beta)} \le P(x^N, y^N) \le 2^{-N(H(X,Y)-\beta)},$$

а вероятностная мера множества типичных пар заключена в интервале

$$(1 - \frac{\sigma^2}{N\beta^2}) \le P(J_\beta) \le 1.$$

Чтобы оценить сверху мощность множества J_β, достаточно поделить верхнюю оценку его полной вероятности (единицу) на нижнюю оценку вероятности элемента – $2^{-N(H(X,Y)+\beta)}$.

6.10. Граница для вероятности независимого выбора типичной пары 38-2

Пусть $P(X,Y)$ – совместное распределение вероятностей с энтропией $H(X,Y)$, а $P(X)$, $P(Y)$ – соответствующие маргинальные распределения с энтропиями $H(X)$, $H(Y)$. Рассмотрим порожденные ими распределения вероятностей N-блоков

$$X^N = \{x^N = (x_1, x_2, \ldots, x_N);\, x_j \in X\} \quad Y^N = \{y^N = (y_1, y_2, \ldots, y_N);\, y_j \in Y\}:$$

$$P(x^N, y^N) = \prod_j P(x_j, y_j); \quad P(x^N) = \prod_j P(x_j); \quad P(y^N) = \prod_j P(y_j).$$

Введем множество J_β совместно типичных пар:

$$J_\beta = \Big\{(x^N, y^N) : \Big|\frac{1}{N}\log\frac{1}{P(x^N, y^N)} - H(X,Y)\Big| \le \beta\Big\}$$

Пусть блоки x^N, y^N выбраны независимо согласно маргинальным распределениям $P(X^N)$, $P(Y^N)$. Показать, что вероятность того, что пара (x^N, y^N) окажется совместно типичной не превышает

$$P\Big((x^N, y^N) \in J_\beta\Big) \le 2^{-N(I(X,Y)-3\beta)},$$

где $I(X,Y)$ – взаимная информация между X и Y.

Решение: Имеем

$$P\Big((x^N, y^N) \in J_\beta\Big) = \sum_{(x^N, y^N) \in J_\beta} P(x_N)P(y_N).$$

Но,

$$P(x_N) \le 2^{-N(H(X)-\beta)}; \quad P(y_N) \le 2^{-N(H(Y)-\beta)}.$$

Поэтому

$$P\Big((x^N, y^N) \in J_\beta\Big) \le \Big|J_\beta\Big| 2^{-N(H(X)+H(Y)-2\beta)} \le 2^{N(H(X,Y)+\beta)} 2^{-N(H(X)+H(Y)-2\beta)}.$$

6.11. Лемма о выбрасывании

Пусть для некоторого канала предложен код $C = \{c^N = (c_1, \ldots, c_N)\}$ длины N и мощности $M = |C|$ и схема декодирования, такие что средняя вероятность P_e ошибки на блок не превышает ϵ:

$$P_e = \sum_{c^N \ne \tilde{c}^N} P(c^N, \tilde{c}^N) < \epsilon.$$

Покажите, что выбросив не более половины из M кодовых слов можно добиться того, чтобы максимальная вероятность ошибки

$$P_{max} = \max_{c^N} \max_{\tilde{c}^N \ne c^N} P(c^N, \tilde{c}^N)$$

не превышала 2ϵ. К какой потере скорости кода приведет такое выбрасывание ?

Решение: Слово c^N назовем плохим (подлежащим выбрасыванию), если для него

$$\max_{\tilde{c}^N \ne c^N} P(c^N, \tilde{c}^N) > 2\epsilon$$

Ясно, что если плохих слов более половины, то средняя вероятность ошибки превышает ϵ. Выбросив все плохие слова (не более половины), получим код с $P_{max} < 2\epsilon$. При выбрасывании половины слов скорость кода снизится незначительно – с $R = \frac{\log M}{N}$ до $R_{ex} = \frac{\log M/2}{N} = \frac{(\log M - 1)}{N} = R - \frac{1}{N}$.

6.12. Границы для вероятности ошибки на бит 33-2

Пусть кодер канала отображает равновероятные двоичные K-блоки на 2^K кодовых слов c. Декодер выносит решения $\tilde{c}$ относительно переданных слов со средней вероятностью ошибки на слово, равной $P_e = \sum_{\tilde{c}\neq c} P(c, \tilde{c})$. Показать, что для средней вероятности ошибки в переданном P_b бите имеют место границы:

$$\frac{P_e}{K} \leq P_b \leq P_e.$$

Решение: Пусть P_k – вероятность ошибки в k-бите, $P_b = \frac{1}{K}\sum_k P_k = \frac{N_b}{K}$, где $N_b = \sum_k P_k$ – матожидание числа ошибочных битов. Пусть $d_b(\tilde{c}, c)$ –число ошибочных битов, возникающих при вынесении ошибочного решения $\tilde{c}$ вместо c. Ясно, что

$$N_b = \sum_{\tilde{c}\neq c} d_b(\tilde{c}, c) P(c, \tilde{c}).$$

но $1 \leq d_b(\tilde{c}, c) \leq K$. Так что $P_e \leq N_b \leq KP_e$. Осталось поделить все на K.

6.13. Пропускная способность двоичного симметричного канала 34-2

Найти пропускную способность C двоичного симметричного канала с входом $X = 0, 1$, выходом $Y = 0, 1$ и вероятностями ошибки $P(y = 1|x = 0) = P(y = 0|x = 1) = p$. Построить график зависимости $C(p)$.

Решение:

$$C = \max_{P(X)} I(X, Y) = \max_{P(X)} [H(Y) - H(Y|X)].$$

Канал симметричен по выходу, поскольку $H(Y|0) = H(Y|1) = -p\log p - (1-p)log(1-p) = h(p)$, так что

$$C = \max_{P(X)} I(X, Y) = \max_{P(X)} [H(Y)] - h(p).$$

Равный 1 максимум $H(Y)$ достигается при равномерном распределении выходов, которое обеспечивается равномерным распределением входов, поскольку

$$P(y) = \sum_x P(y|x)P(x) = \frac{1}{2}\sum_x P(y|x) = \frac{1}{2}$$

не зависит от y. Так что $C = 1 - h(p)$.

6.14. Пропускная способность q-ичного симметричного канала 35-2

Найти пропускную способность C q-ичного симметричного канала со входом $X = 1, 2, \ldots, q$, выходом $Y = 1, 2, \ldots, q$ и вероятностями $P(y = x) = 1 - p$ $P(y \neq x) = \frac{p}{q-1}$. Построить график зависимости $C(p)$. При каком значении p пропускная способность обращается в нуль ?

Решение: Все условные энтропии $H(Y|x) = h_q(p) = h(p) + p\log(q-1)$ одинаковы, а максимум $H(Y) = \log q$ достигается на равномерном распределении входов. Поэтому $C = \log q - h_q(p) = \log q - h(p) - p\log(q-1)$. Пропускная способность обращается в нуль, когда $1 - p = \frac{p}{1-q}$ (все выходы равновероятны), то есть, при $p = \frac{q-1}{q}$.

6.15. Сумасшедшая пишущая машинка 36-2

Найти пропускную способность C q-ичной сумасшедшей пишущей машинки – канала с одинаковыми входным и выходным алфавитами $X = Y = (0, 1, \dots, q-1)$ и такого, что каждый данный символ s переходит в себя с вероятностью $1 - 2p$, а с равными вероятностями p отображается на соседние символы $(s-1) \mod q$ и $(s+1) \mod q$. Для частного случая $p = \frac{1}{3}$ и $q = 3^m$ предложить схему кодирования, достигающую пропускной способности.

Решение: Все условные энтропии $H(Y|x)$ одинаковы, а равномерное распределение входов дает равномерное распределение выходов – все суммы $\sum_x P(y|x)$ одинаковы. Поэтому

$$C = H(Y) - H(Y|x) = \log q + (1-2p)\log(1-2p) + 2p\log p.$$

При $p = 1/3$, $q = 3^m$ получается: $C = (m-1)\log 3$. Эта пропускная способность достигается, если использовать только треть символов – 3^{m-1} из 3^m имеющихся, то есть нажимать клавиши равновероятно, но с шагом 3 по кругу.

6.16. Пропускная способность канала со стираниями 39-2

Найти пропускную способность C двоичного симметричного канала со стираниями: вход $X = 0, 1$, выход $Y = 0, 1, z$, $p(z|0) = p(z|1) = p$, $p(0|0) = p(1|1) = 1 - p$. Построить график зависимости $C(p)$.

Решение:

$$C = \max_{P(X)} I(X,Y) = \max_{P(X)}[H(Y) - H(Y|X)].$$

Канал симметричен по выходу, поскольку $H(Y|0) = H(Y|1) = h(p)$. Поэтому

$$C = \max_{P(X)} I(X,Y) = \max_{P(X)}[H(Y)] - h(p).$$

Входы эквивалентны. Поэтому максимум $H(Y)$ достигается при равновероятных входах. При этом распределение вероятностей выходов имеет вид $P(0) = P(1) = \frac{1-p}{2}$, $P(z) = p$. Энтропия этого распределения составляет

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

Так что $C = H(Y) - h(p) = 1 - p$.

Или иначе: $C = H(X) - H(X|Y)$, но $H(X|Y) = P(Y = z)$. Но вероятность наблюдениия стертого символа на выходе не зависит от распределения на входе и составляет p. Поэтому $C = \max_{P(X)} H(X) - p = 1 - p$.

6.17. Пропускная способность Z-канала 40-2

Найти выражение для пропускной способности $C(p)$ Z-канала с двоичным входом $X = \{0, 1\}$, двоичным выходом $Y = \{0, 1\}$ и матрицей условных вероятностей $P(y = 0|x = 0) = 1$, $P(y = 1|x = 1) = p$, $P(y = 0|x = 1) = 1 - p$. Найти численное значение пропускной способности при $p = 1/2$. Каковы предельные значения $C(p)$ при $p \to 0$ и $p \to 1$.

Решение:

Канал несимметричен, так что его пропускная способность не достигается на равномерном распределении входов. Пусть $q = P(x = 1)$ и $1 - q = P(x = 0)$. Тогда $P(y = 1) = qp$, $P(y = 0) = q(1-p) + (1-q) = 1 - qp$, так что

$$I(X|Y) = H(Y) - H(Y|X) = h(qp) - qh(p),$$

где $h(x)$ – двоичная энтропия. Приравняв к нулю производную по q найдем экстремальное значение

$$q = \frac{1/p}{1 + 2^{h(p)/p}}.$$

В итоге для пропускной способности получается: $C(p) = \log(1 + 2^{\frac{h(p)}{p}}) - \frac{h(p)}{p}$. При $p = 1/2$ $q = 2/5$, $C = \log 5 - 2 > 0$. $\lim_{p\to 0} C(p) = 0$, $\lim_{p\to 1} C(p) = 1$,

6.18. Пропускная способность параллельного соединения независимых каналов - общий случай 1-3

Пусть два независимые канала со входами X_1, X_2 и выходами Y_1, Y_2 соединены параллельно, образуя векторный канал с матрицей условных вероятностей $P(Y_1Y_2|X_1X_2) = P(Y_1|X_1)P(Y_2|X_2)$. Покажите, что пропускная способность параллельного соединения равна сумме пропускных способностей каналов.

Решение:

$$C = \max_{P(X_1,X_2)} [H(Y_1Y_2) - H(Y_1Y_2|X_1X_2)].$$

Но, в силу независимости подканалов, $H(Y_1Y_2|X_1X_2) = H(Y_1|X_1) + H(Y_2|X_2)$. Далее, энтропия $H(Y_1Y_2) \le H(Y_1) + H(Y_2)$ максимальна, когда случайные величины Y_1, Y_2 независимы. Но это заведомо имеет место при $P(X_1X_2) = P(X_1)P(X_2)$. Так что

$$C = \max_{P(X_1),P(X_2)} [H(Y_1) - H(Y_1|X_1) + H(Y_2) - H(Y_2|X_2)] = C_1 + C_2.$$

6.19. Пропускная способность параллельного соединения независимых каналов 2-3

Найти пропускную способность параллельного соединения пары двоичных симметричных каналов с вероятностями искажения символа p и q – канала с векторным входом (X_1, X_2), $X_1, X_2 = \{0, 1\}$, векторным выходом (Y_1, Y_2), $Y_1, Y_2 = \{0, 1\}$ и вероятностями ошибки p в субканале $X_1 \to Y_1$ и q – в субканале $X_2 \to Y_2$.

Решение: Условная энтропия $H(Y_1, Y_2|x_1, x_2)$ одинакова для всех пар входов и составляет $h(p) + h(q)$. Равномерное распределение входов дает равномерное распределение выходов с энтропией $H(Y_1, Y_2) = 2$. Поэтому $C = 2 - h(p) - h(q) = C(p) + C(q)$, где $C(p) - 1 - h(p)$, $C(q) - 1 - h(q)$.

6.20. Пропускная способность параллельных каналов с общим входом 3-3

Найти пропускную способность параллельного соединения пары двоичных симметричных каналов с объединенным входом $X = \{0, 1\}$, векторным выходом (Y_1, Y_2) $Y_1, Y_2 = \{0, 1\}$ и вероятностями p и q искажения в субканалах $X \to Y_1$ и $X \to Y_2$. Учесть, что пропускная способность достигается на равномерном распределении вероятностей входов. Какой окажется эта пропускная способность при $q = 0$, $q = 1/2$, $q = 1$.

Решение: Наборы условных вероятностей выходов при передаче нуля и единицы имеют вид

$$P(00|0) = (1-p)(1-q); \quad P(10|0) = p(1-q); \quad P(01|0) = (1-p)q; \quad P(11|0) = pq,$$

$$P(00|1) = pq; \quad P(10|1) = (1-p)q; \quad P(01|1) = p(1-q); \quad P(11|1) = (1-p)(1-q).$$

Их энтропии одинаковы и ожидаемо составляют $h(p)+h(q)$. При равномерном распределении входов

$$P(Y_1Y_2=00)=P(Y_1Y_2=11)=\frac{1-p-q+2pq}{2},$$

$$P(Y_1Y_2=10)=P(Y_1Y_2=01)=\frac{p+q-2pq}{2}.$$

Так что $H(Y_1,Y_2)=1+h(p+q-2pq)$ и $C(p,q)=1+h(p+q-2pq)-h(p)-h(q)$. Элементарная проверка дает: $C(p,0)=C(p,1)=1$, $C(p,1/2)=1-h(p)$.

6.21. Дифференциальная энтропия одномерной гауссовской плотности 4-3

Найти дифференциальную энтропию $H=\int_x g(x)\log\frac{1}{g(x)}dx$ гауссовской плотности вероятностей

$$g(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

с дисперсией σ^2 и средним значением μ. Показать, что

$$H=\int_x \rho(x)\log\frac{1}{g(x)}dx, \quad \text{если} \quad \int_x x^2\rho(x)dx=\sigma^2.$$

Решение: Простые выкладки дают ответ: $H=\log\sqrt{2\pi e\sigma^2}=\frac{1}{2}\log 2\pi e\sigma^2$. Независимость от $\rho(x)$ видна по ходу вычислений.

6.22. Граница для информационной дивергенции 5-3

Пусть $p(x), g(x)$ – две плотности вероятностей. Показать, что

$$\int p(x)\log\frac{p(x)}{g(x)}dx\geq 0$$

с равенством при $p(x)=g(x)$.

Решение:

Применим неравенство Йенсена для выпуклой вниз функции $-\log(x)$:

$$\int p(x)\log\frac{p(x)}{g(x)}dx=E_p\Big[-\log\frac{g(x)}{p(x)}\Big]\geq-\log E_p\Big[\frac{g(x)}{p(x)}\Big]=-\log 1=0.$$

6.23. Дифференциальная энтропия двумерной гауссовской плотности 6-3

Найти дифференциальную энтропию двумерной гауссовской плотности вероятностей

$$g(z)=\frac{1}{2\pi\sigma^2}e^{-\frac{|z|^2}{2\sigma^2}}, \quad z=x+jy, |z^2|=x^2+y^2,$$

с дисперсией $2\sigma^2$.

Решение: Простые выкладки дают ответ: $H=\log 2\pi e\sigma^2$.

6.24. Экстремальность гауссовской плотности 7-3

Показать, что в классе плотностей вероятностей $\varrho(x)$ с нулевым средним и заданной дисперсией $\sigma^2 = \int x^2 \varrho(x) dx$ гауссовская плотность

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

обладает максимальной энтропией.

Решение:

Достаточно показать, что

$$\int \rho(x) \log \frac{1}{\rho(x)} dx \le \int \rho(x) \log \frac{1}{g(x)} dx = \int g(x) \log \frac{1}{g(x)} = \log \sqrt{2\pi e \sigma^2}.$$

Но это следует из универсальной границы

$$\int \rho(x) \log \frac{\rho(x)}{g(x)} dx \ge 0$$

для информационной дивергенции.

Иначе, максимизируем по $\rho(x)$ энтропию при фиксированной дисперсии. Варьирование по ρ лагранжиана

$$L(\rho) = -\int \rho(x) \ln \rho(x) dx - \lambda \int x^2 \rho(x) dx - \gamma \int \rho(x) dx$$

дает

$$-\ln\rho - 1 - \lambda x^2 - \gamma = 0; \quad \Rightarrow \quad \rho(x) = e^{-(1+\gamma)} e^{-\lambda x^2}.$$

Значения постоянных λ и γ дают ограничения на дисперсию и условие нормировки на единицу.

6.25. Вероятность ошибки в двоичном гауссовском канале - максимум правдоподобия 8-3

Пусть бит (0 или 1) передается по каналу противоположными сигналами $x_0(t) = +cp(t)$, $x_1(t) = -cp(t)$, где $\int p^2(t)dt = 1$, а $E_c = c^2 = \int x_{0,1}^2(t)dt$ – энергия сигнала. Принятия реализация $y(t) = x_{0,1}(t) + n(t)$ отличается добавлением белого гауссовского шума $n(t)$. Согласованный фильтр приемника вычисляет проекцию $y(t)$ на опорный импульс $p(t)$. Результатом

$$y = \int y(t)p(t)dt = \pm c \int p^2(t)dt + \int n(t)p(t)dt = \pm c + w$$

оказывается отсчет y, равный $\pm c$ плюс случайная шумовая добавка w с дисперсией σ^2. Решения относительно переданного бита выносятся по максимуму правдоподобия. Найти зависимость средней вероятности ошибочного P_e от отношения сигнал/шум $\mu^2 = \frac{c^2}{\sigma^2} = \frac{E_c}{\sigma^2}$.

Решение:

Шум w – гауссовский с плотностью

$$\rho(w) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{w^2}{2\sigma^2}}.$$

Условные плотности распределения отсчета y имеют вид

$$\varrho(y/0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-c)^2}{2\sigma^2}}; \quad \varrho(y/1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+c)^2}{2\sigma^2}};$$

Решение максимума правдоподобия выносится по знаку y. Вероятность ошибочного решения при передаче бита 1 (символа $-c$), она же средняя вероятность ошибки, составляет

$$P_e = \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty e^{-\frac{(y+c)^2}{2\sigma^2}} dy = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{(x+\mu)^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_\mu^\infty e^{-\frac{x^2}{2}} dx = Q(\mu),$$

где $Q(x)$ - Q-функция

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_\mu^\infty e^{-\frac{x^2}{2}} dx.$$

Таким образом, $P_e = Q(\mu) = Q\left(\sqrt{\frac{E_c}{\sigma^2}}\right)$.

6.26. Вероятность ошибки в двоичном гауссовском канале - максимум апостериорной вероятности 10-3

Пусть бит (0 или 1) передается по каналу противоположными сигналами $x_0(t) = +cp(t)$, $x_1(t) = -cp(t)$, где $\int p^2(t)dt = 1$, а $E_c = c^2 = \int x_{0,1}^2(t)dt$ – энергия сигнала. Принятия реализация $y(t) = x_{0,1}(t) + n(t)$ отличается добавлением белого гауссовского шума $n(t)$. Согласованный фильтр приемника вычисляет проекцию $y(t)$ на опорный импульс $p(t)$. Результатом

$$y = \int y(t)p(t)dt = \pm c \int p^2(t)dt + \int n(t)p(t)dt = \pm c + w$$

оказывается отсчет y, равный $\pm c$ плюс случайная шумовая добавка w с дисперсией σ^2. Решения относительно переданного бита выносятся по максимуму апостериорной вероятности с априороной гипотезой о том, что P(1) = q. Найти зависимость средней вероятности ошибочного P_e от отношения сигнал/шум $\mu^2 = \frac{c^2}{\sigma^2} = \frac{E_c}{\sigma^2}$. Сколь малым должно быть q, чтобы демодулятр максимума апостериорной вероятности принимал равновероятные решения при передаче единицы.

Решение:

Нормируем выход на σ, положив $x = \frac{y}{\sigma}$. Придем к двоичному каналу с условными плотностями распределения выхода x

$$\varrho(x|0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}; \quad \varrho(x|1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+\mu)^2}{2}};$$

Для логарифма отношения правдоподобия апостериорных плотностей вероятностей найдем:

$$\lambda = \ln \frac{P(0|x)}{P(1|x)} = \ln \frac{\rho(x|0)(1-q)}{\rho(x|1)q} = 2\mu x + \lambda_0; \quad \lambda_0 = \ln \frac{1-q}{q}.$$

Решения выносятся по результату сравнения x с порогом $-\frac{\lambda_0}{2\mu}$. При $q < 1/2$ этот порог смещен на отрицательную полуось x. Вероятность ошибки при передаче 1 составляет $P_e(1) = Q(\mu - \frac{\lambda_0}{2\mu})$, а при передаче нуля – $P_e(0) = Q(\mu + \frac{\lambda_0}{2\mu})$, так что средняя вероятность ошибки составляет

$$P_e = qQ(\mu - \frac{\lambda_0}{2\mu}) + (1-q)Q(\mu + \frac{\lambda_0}{2\mu})$$

При $\lambda_0 = \ln\frac{1-q}{q} = 2\mu^2$ $(q = \frac{1}{1+e^{2\mu^2}})$ единичный бит демодулируется случайно – с вероятностью $1/2$. Вероятность ошибки демодуляции нулевого бита составляет при этом $Q(2\mu) \ll Q(\mu)$.

6.27. Пропускная способность вещественного непрерывного гауссовского канала 9-3

Найти пропускную способность вещественного непрерывного гауссовского канала $y = x+w$ со средней энергией передаваемого вещественного символа $E[x^2] = c^2 = E$ и дисперсией шума $E[w^2] = \sigma^2 = \mathcal{N}_0/2$. Выразить ее через энергию на бит $E_b = E/R$ и одностороннюю спектральную плотность шума $\mathcal{N}_0$. Показать, что надежная передача данных возможно только при $\frac{E_b}{\mathcal{N}_0} > \ln 2$.

Решение:

Введем отношение сигнал/шум $\mu^2 = \frac{c^2}{\sigma^2}$ и нормируем все на σ. Получим канал $y' = x' + w'$ с дисперсиями $E[x'^2] = E[x^2/\sigma^2] = \mu^2$, $E[w'^2] = E[w^2/\sigma^2] = 1$, $E[y'^2] = E[x'^2] + E[w'^2] = \mu^2 + 1$.

Энтропия условной плотности $H(Y|x)$ одинакова для всех x, равна энтропии гауссовской плотности $\frac{1}{\sqrt{2\pi}}e^{-\frac{(y-x)^2}{2}}$ и составляет $\log\sqrt{2\pi e}$. Осталось максимизировать энтропию $H(Y)$ по плотностям распределения входов с фиксированной дисперсией μ^2.

В классе плотностей $\rho(y)$ с фиксированной дисперсией $\mu^2 + 1$ максимальной энтропией $\log\sqrt{2\pi e(\mu^2+1)}$ обладает гауссовское. Но при гауссовском распределении входов оно получается автоматически. Поэтому пропускная способность достигается на гауссовском распределении входов:

$$C = H(Y) - H(Y|x) = \log\sqrt{2\pi e(1+\mu^2)} - \log\sqrt{2\pi e} = \frac{1}{2}\log(1+\mu^2) = \frac{1}{2}\log(1+\frac{E}{\sigma^2}).$$

Для скорости надежной передачи $R < C$ получаются границы

$$R < \frac{1}{2}\log(1+\mu^2); \quad \mu^2 = \frac{E}{\sigma^2} = \frac{2RE_b}{\mathcal{N}_0} > 2^{2R} - 1$$

или

$$\frac{E_b}{\mathcal{N}_0} > \frac{2^{2R}-1}{2R} > \ln 2,$$

поскольку $\lim_{x\to 0}\frac{2^x-1}{x} = \ln 2$.

6.28. Пропускная способность комплексного непрерывного гауссовского канала 11-3

Найти пропускную способность непрерывного комплексного канала $y = x + w$ со средней энергией передаваемого вещественного символа $E[|x|^2] = c^2 = E$ и дисперсией шума $E[|w|^2] = 2\sigma^2 = \mathcal{N}_0$. Выразить ее через энергию на бит $E_b = E/R$ и одностороннюю спектральную плотность шума $\mathcal{N}_0$. Показать, что надежная передача данных возможно только при $\frac{E_b}{\mathcal{N}_0} > \ln 2$.

Решение:

Введем отношение сигнал/шум $\mu^2 = \frac{c^2}{2\sigma^2}$ и нормируем все на σ. Получим канал $y' = x' + w'$ с дисперсиями $E[x'^2] = E[x^2/\sigma^2] = 2\mu^2$, $E[w'^2] = E[w^2/\sigma^2] = 2$, $E[y'^2] = E[x'^2] + E[w'^2] = 2\mu^2 + 2$.

Энтропия условной плотности $H(Y|x)$ одинакова для всех x, равна энтропии гауссовской плотности $\frac{1}{2\pi}e^{-\frac{|y-x|^2}{2}}$ и составляет $\log 2\pi e$. Осталось максимизировать энтропию $H(Y)$ по плотностям распределения входов с фиксированной дисперсией μ^2.

В классе плотностей $\rho(y)$ с фиксированной дисперсией $2 + 2\mu^2$ максимальной энтропией $\log 2\pi e(2 + 2\mu^2)$ обладает гауссовское. Но при гауссовском распределении входов оно получается автоматически. Поэтому пропускная способность достигается на гауссовском распределении входов:

$$C = H(Y) - H(Y|x) = \log 2\pi e(2\mu^2 + 2) - \log(4\pi e) = \log(1 + \mu^2) = \log(1 + \frac{E}{2\sigma^2}).$$

Для скорости надежной передачи $R < C$ получаются границы

$$R < \log(1 + \mu^2); \quad \mu^2 = \frac{E}{2\sigma^2} = \frac{RE_b}{\mathcal{N}_0} > 2^R - 1$$

или

$$\frac{E_b}{\mathcal{N}_0} > \frac{2^{2R} - 1}{2R} > \ln 2,$$

поскольку $\lim_{x\to 0} \frac{2^x - 1}{x} = \ln 2$.

6.29. Параллельные гауссовские каналы 12-3

Найти пропускную способность системы из N параллельных вещественных непрерывных гауссовских каналов $y_n = x_n + w_n$, $n = 1..N$ со средней энергией передаваемого вещественного символа $E[x_n^2] = E_n$, дисперсией шума $E[w_n^2] = \sigma_n^2$. Какое распределение энергий между каналами с фиксированной полной энергией $E_0 = \sum_n E_n$ максимизирует эту пропускную способность? Как распределить энергию E_0, когда все дисперсии одинаковы? Какой окажется при этом пропускная способность?

Решение:

Пропускная способность параллельного соединения независимых каналов равна сумме пропускных способностей:

$$C = \frac{\log e}{2} \sum_n \ln(1 + \frac{E_n}{\sigma_n^2}).$$

Требуется максимизировать эту сумму по E_n при ограничении $E_0 = \sum_n E_n$. Приравнивание к нулю производной лагранжиана

$$L = \sum_n \ln(1 + \frac{E_n}{\sigma_n^2}) - \frac{1}{\lambda} \sum_n E_n$$

дает $E_n = \lambda - \sigma_n^2$. Ограничение $E_0 = \sum_n E_n = N\lambda - \sum_n \sigma_n^2$ определяет значение $\lambda = \frac{E_0 + \sum_n \sigma_n^2}{N}$, что и дает ответ

$$E_n = \frac{E_0 + \sum_n \sigma_n^2}{N} - \sigma_n^2$$

При одинаковых $\sigma_n^2 = \sigma^2$ энергия распределяется поровну – $E_n = \frac{E}{N}$ и

$$C = \frac{N}{2}(1 + \frac{E_0}{N\sigma^2}).$$

6.30. Вещественные и комплексные гауссовские каналы

Найти пропускную способность пары параллельных вещественных непрерывных гауссовских каналов $y_n = x_n + w_n$, $n = 1, 2$ со средней энергией передаваемого вещественного символа $E[x_n^2] = E_n$, дисперсией шума $E[w_n^2] = \sigma_n^2$. Какое распределение энергий между каналами при фиксированной полной энергии $E_0 = E_1 + E_2$ максимизирует эту пропускную способность? Как распределить энергию E_0, когда обе дисперсии одинаковы? Показать, что при одинаковых дисперсиях пропускная способность параллельного соединения равна пропускной способности комплексного гауссовского канала.

Решение:

Согласно предыдущей задаче, оптимальное распределение энергий имеет вид

$$E_1 = \frac{E_0 + \sigma_2^2 - \sigma_1^2}{2}; \quad E_2 = \frac{E_0 + \sigma_1^2 - \sigma_2^2}{2}.$$

(Большую долю энергии следует направить в плохой канал). При $\sigma_1^2 = \sigma_2^2 = \sigma^2$ энергия делится поровну, а для пропускной способности получается

$$C = \log(1 + \frac{E_0}{2\sigma^2}) = \log(1 + \mu^2),$$

что как раз совпадает с пропускной способностью комплексного канала.

6.31. К параллельному соединению каналов 13-3

Пусть имеется пара непрерывных вещественных гауссовских каналов с одинаковой дисперсией шума σ^2. Для передачи символа выделена фиксированная энергия E. Что лучше в плане пропускной способности – вложить всю эту энергию в один канал, или распределить ее между двумя каналами поровну? Оценить выигрыш в пропускной способности. Как этот выигрыш зависит от отношения сигнал/шум $\mu^2 = \frac{E}{\sigma^2}$?

Решение: При передаче по одному каналу $C_0 = \frac{1}{2}\log(1 + \frac{E}{\sigma^2})$. При использовании двух каналов – $C_{\#} = \log(1 + \frac{E}{2\sigma^2})$. Ясно, что $C_{\#} > C_0$ в той мере, в какой

$$(1 + \frac{E}{2\sigma^2})^2 = 1 + \frac{E}{\sigma^2} + \frac{E^2}{4\sigma^4} > 1 + \frac{E}{\sigma^2}.$$

Вообще,

$$2(C_{\#} - C_0) = \log\Big(1 + \frac{\mu^2}{4}\frac{\mu^2}{1+\mu^2}\Big).$$

Конкретно, при $\mu^2 = \frac{E}{\sigma^2} = 1$ $C_0 = \frac{1}{2}$, а $C_{\#} = \log 3 - 1 = 0.585$.

6.32. Предельная пропускная способность системы параллельных каналов

Пусть для передачи символа выделена фиксированная энергия E_0. Если всю ее вложить в один гаусовский канал с дисперсией шума $\sigma^2 = \mathcal{N}_0/2$, получится пропускная способность $C = \frac{1}{2}\log(1 + \frac{2E_0}{\mathcal{N}_0})$. Какой пропускной способности можно достичь, равномерно распределив энергию E_0 между $N \to \infty$ одинаковым гауссовскими каналами ?

Решение:

$$C_N = \frac{N}{2}\log(1 + \frac{1}{N}\frac{2E_0}{\mathcal{N}_0}) \to \frac{E_0}{\mathcal{N}_0}\log e.$$

6.33. (Предельная скорость передачи по радиоканалу 1 14-3

Пусть отношение сигнал/шум в комплексном радиоканале составляет $\mu^2 = \frac{E_s}{2\sigma^2} = \frac{E_s}{\mathcal{N}_0} = 7$. Какова при этом предельная скорость R надежной передачи данных в битах на измерение? Какова реальная скорость передачи R_b в битах в секунду, если полоса канала составляет $F = 1MHz$. Каково при этом отношение $\frac{E_b}{\mathcal{N}_0}$? Насколько оно удалено от шенноновского предела $\frac{E_b}{\mathcal{N}_0} > \ln 2$?

Решение:

Согласно границе Шеннона $R < \log(1+\mu^2) = \log(1+7) = 3$ битам на символ. При предельной спектральной эффективности системы сигналов $\rho = 1$ символов в секунду на Герц полосы скорость передачи составит $R_{bh} = \rho R = 3$ бита в секунду на герц полосы или $R_b = \rho RF = 3.10^6$ битов в секунду. Отношение $\frac{E_b}{\mathcal{N}_0} = \frac{E_s}{R\mathcal{N}_0}$ составляет $7/3 = 2.33 > \ln 2 = 0.69$

6.34. Предельная скорость передачи по радиоканалу 2 15-3

Пусть отношение сигнал/шум в комплексном радиоканале $\mu^2 = \frac{E_s}{2\sigma^2} = \frac{E_s}{\mathcal{N}_0}$ составляет 0.0718. Предельная скорость R надежной передачи данных в битах на измерение составляет при этом $R = \log(1+\mu^2) = 0.1$. Какова реальная предельная скорость передачи R_b в битах в секунду, если полоса канала составляет $F = 1MHz$? Каково отношение $\frac{E_b}{\mathcal{N}_0}$? Насколько оно удалено от шенноновского предела $\frac{E_b}{\mathcal{N}_0} > \ln 2$

Решение:

При предельной спектральной эффективности системы сигналов $\rho = 1$ символов в секунду на Герц полосы скорость передачи составит $R_{bh} = \rho R = 0.1$ бита в секунду на герц полосы или $R_b = \rho RF = 3.10^5$ битов в секунду. Отношение $\frac{E_b}{\mathcal{N}_0} = \frac{E_s}{R\mathcal{N}_0}$ составляет $0.0718/0.1 = 0.718 > \ln 2 = 0.69$

6.35. Предельная скорость передачи по радиоканалу 3 17-3

Пусть отношение сигнал/шум в вещественном канале составляет $\mu^2 = \frac{E_s}{\sigma^2} = \frac{2E_s}{\mathcal{N}_0} = 1$. Какова при этом предельная скорость R надежной передачи данных в битах на измерение. Какова реальная скорость передачи R_b в битах в секунду, если полоса канала составляет $F = 1MHz$. Каково при этом отношение $\frac{E_b}{\mathcal{N}_0}$. Насколько оно удалено от шенноновского предела $\frac{E_b}{\mathcal{N}_0} > \ln 2$

Решение:

Согласно границе Шеннона $R < \frac{1}{2}\log(1+\mu^2) = \frac{1}{2}\log(1+1) = 1/2$ бита на символ. При предельной спектральной эффективности системы сигналов $\rho = 1$ символов в секунду на Герц полосы скорость передачи составит $R_{bh} = \rho R = 1/2$ бита в секунду на герц полосы или $R_b = \rho RF = 5.10^5$ битов в секунду. Отношение $\frac{E_b}{\mathcal{N}_0} = \frac{E_s}{2R(\mathcal{N}_0/2)}$ составляет $1 > \ln 2 = 0.69$

6.36. Пропускная способность двоичного гауссовского канала - жесткие решения 16-3

Пусть бит (0 или 1) передается по каналу противоположными сигналами $x_0(t) = +cp(t)$, $x_1(t) = -cp(t)$, где $\int p^2(t)dt = 1$, а $E_c = c^2 = \int x_{0,1}^2(t)dt$ – энергия сигнала. Принятия реализация $y(t) = x_{0,1}(t) + n(t)$ отличается добавлением белого гауссовского шума $n(t)$. Согласованный фильтр приемника вычисляет проекцию $y(t)$ на опорный импульс $p(t)$. Результатом

$$y = \int y(t)p(t)dt = \pm c \int p^2(t)dt + \int n(t)p(t)dt = \pm c + w$$

оказывается отсчет y, равный $\pm c$ плюс случайная шумовая добавка w с дисперсией σ^2. Жесткие решения относительно переданного бита выносятся по максимуму правдоподобия. Найти зависимость $C_h(\mu)$ пропускной способности этого канала отношения сигнал/шум $\mu^2 = \frac{c^2}{\sigma^2} = \frac{E_c}{\sigma^2}$.

Решение: Имеем двоичный симметричный канал с вероятностью ошибки $P_e = Q(\mu)$. Так что

$$C_h(\mu) = 1 - h(Q(\mu)).$$

6.37. Пропускная способность двоичного гауссовского канала - мягкие решения 19-3

Пусть бит (0 или 1) передается по каналу противоположными сигналами $x_0(t) = +cp(t)$, $x_1(t) = -cp(t)$, где $\int p^2(t)dt = 1$, а $E_c = c^2 = \int x_{0,1}^2(t)dt$ – энергия сигнала. Принятия реализация $y(t) = x_{0,1}(t) + n(t)$ отличается добавлением белого гауссовского шума $n(t)$. Согласованный фильтр приемника вычисляет проекцию $y(t)$ на опорный импульс $p(t)$. Результатом

$$y = \int y(t)p(t)dt = \pm c \int p^2(t)dt + \int n(t)p(t)dt = \pm c + w$$

оказывается отсчет y, равный $\pm c$ плюс случайная шумовая добавка w с дисперсией σ^2. Решения относительно переданного бита по выносятся наблюдению y по максимуму правдоподобия. Найти выражение для $C_s(\mu)$ – зависимости пропускной способности этого канала от отношения сигнал/шум $\mu^2 = \frac{c^2}{\sigma^2} = \frac{E_c}{\sigma^2}$.

Решение: Нормируем выход на σ, положив $x = \frac{y}{\sigma}$. Придем к двоичному каналу с условными плотностями распределения выхода x

$$\varrho(x|0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}}; \quad \varrho(x|1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+\mu)^2}{2}};$$

Энтропии этих распределений одинаковы и составляют $\log\sqrt{2\pi e}$. Пропускная способность достигается на равновероятных входах – симметрия. Остается вычислить энтропию безусловного распределения выхода

$$\varrho(x) = \frac{1}{2\sqrt{2\pi}} \left[e^{-\frac{(x-\mu)^2}{2}} + e^{-\frac{(x+\mu)^2}{2}} \right]$$

Несложные вычисления с учетом тождества (замена x на $-x$)

$$\int \log[e^{-\frac{(x-\mu)^2}{2}} + e^{-\frac{(x+\mu)^2}{2}}\Big] e^{-\frac{(x-\mu)^2}{2}} dx = \int \log[e^{-\frac{(x-\mu)^2}{2}} + e^{-\frac{(x+\mu)^2}{2}}\Big] e^{-\frac{(x+\mu)^2}{2}} dx$$

дают

$$C_s(\mu) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2}} \log(1 + e^{-2\mu x}) dx.$$

7. Блочные коды

7.1. К расстоянию Хэмминга 18-3

Показать, что для расстояние Хэмминга $d_H(x, y)$ между двумя n-блоками над q-ичным алфавитом обладает свойствами метрики:

$$0 \le d_H(x, y) \le n,$$

$$d_H(x, y) = 0 \quad \Rightarrow \quad x = y,$$

$$d_H(x, y) \le d_H(x, z) + d_H(z, y).$$

Решение Если в какой-то позиции $x \neq y$, то вклад этой позиции в $d_H(x, y)$ составляет 1, в то время как вклад этой позиции в $d_H(x, z) + d_H(z, y)$ не меньше 1 при любом z.

7.2. Число ошибок 20-3

Показать, что произвольный q-ичный $[n,k,d]_q$-код, $k = \log_q M$ длины n, мощности M с минимальным расстоянием $d = 2t+1$ гарантированно обнаруживает $2t$ ошибок и гарантированно исправляет t ошибок.

Решение

Хэмминговские сферы $V_q(n,t) = \{x : d_H(c,x) \leq t\}$, проведенные вокруг кодовых слов c не перекрываются, поскольку расстояние между словами превышает $2t$.

7.3. Число стираний 21-3

Показать, что произвольный q-ичный $[n,k,d]_q$-код, $k = \log_q M$ длины n, мощности M с минимальным расстоянием d гарантированно обнаруживает $(d-1)$-у ошибку и исправляет $d-1$ стирание.

Решение

Стирание эквивалентно укорочению кода на одну позицию, которое приводит к снижению расстояния максимум на единицу. Укороченные слова остаются различимыми, пока минимальное расстояние не хуже единицы.

7.4. Число ошибок и стираний 22-3

Показать, что произвольный q-ичный $[n,k,d]_q$-код, $k = \log_q M$ длины n, мощности M с минимальным расстоянием $d = 2t+e+1$ гарантированно исправляет e стираний и t ошибок.

Решение После укорочения на e стертых позиций расстояние кода составит не менее $2t+1$.

7.5. Асимптотическая оценка объема сферы 23-3

Пусть $V_q(n,d) = \{x : d_H(c,x) \leq d\}$ – хэмминговская сфера радиуса t. Показать, при $n \to \infty$

$$|V_q(n,d)| \sim 2^{nh(\delta,q)} + o(\delta),$$

где $\delta = \frac{d}{n}$, а $h(\delta,q)) = h(\delta) + \delta \log(q-1)$.

Решение Для числа точке в хэмминговокой сфере справедливы границы:

$$\binom{n}{t}(q-1)^t \leq |V_q(n,t)| = \sum_{k=0}^{t} \binom{n}{k}(q-1)^k \leq (t+1)\binom{n}{t}(q-1)^t$$

Перейдя к логарифмам, поделив все на n и учтя, что $\frac{1}{n}\log\binom{n}{t} \sim h(\delta)$, найдем:

$$h(\delta) + \delta\log(q-1) \leq \frac{1}{n}\log|V_q(n,d)| \leq \frac{1}{n}\log(t+1) + h(\delta) + \delta\log(q-1).$$

Верхняя и нижняя оценки асимптотически совпадают, поскольку

$$\frac{1}{n}\log(t+1) = \frac{1}{n}\log(\delta n + 1) \to 0.$$

7.6. Граница Варшимова-Гильберта 24-3

Доказать существование q-ичного $[n, k, d]_q$-кода с мощностью M, $(k = \log_q M)$, удовлетворяющей границе

$$M \geq \frac{q^n}{|V_q(n, d-1)|}.$$

Вывести отсюда асимптотическую границу

$$R \geq 1 - h_q(\delta),$$

где $R = \frac{\log_q M}{n} = \frac{k}{n}$, $\delta = \frac{d}{n}$, а

$$h_q(\delta) = -\delta \log_q \delta - (1-\delta)\log_q(1-\delta) + \delta \log_q(q-1).$$

Построить график зависимости $R(\delta)$ для $q = 2$. Изучить его поведение с ростом q.

Решение

Применим жадный алгоритм построения случайного кода. Выбрав случайное слово, будем выбрасывать вместе с ним всю окружающую его сферу $V_q(n, d-1)$ радиуса $d-1$. Все оставшиеся слова будут лежать на расстоянии не менее d от выбранных. Пока для числа M уже выбранных слов имеет место оценка $MV_q(n, d-1) < q^n$, во множестве из q^n слов остаются не выброшенные. Процедуру можно продолжить. В результате будет построен код с

$$M \geq \frac{q^n}{|V_q(n, d-1)|}.$$

Асимптотическая граница вытекает из асимптотической оценки для $|V_q(n, d-1)|$:

$$R = \frac{\log_q M}{n} = 1 - \log_q(|V_q(n, d-1)|) = 1 - \frac{\log_2(|V_q(n, d-1)|)}{\log_2(q)} = 1 - \frac{h(\delta, q)}{\log_2(q)} = 1 - h_q(\delta).$$

7.7. Декодирование до границы минимального расстояния 25-3

Покажите, что при декодировании двоичных кодов до границы минимального расстояния (то есть с исправлением на более $t = \frac{d-1}{2}$ ошибок) пропускная способность двоичного симметричного канала не достигается.

Решение

Пропускная способность двоичного симметричного канала с вероятностью ошибки p составляет $C = 1 - h(p)$, в то время как скорость кода, гарантированно исправляющего t ошибок оценивается величиной $R = 1 - h\left(\frac{d}{n}\right) = 1 - h\left(\frac{2t+1}{n}\right)$ (Граница Варшамова Гильберта). При вероятности ошибки p число подлежащих исправлению ошибок составляет порядка $t = pn$, то есть $\frac{t}{n} = p$. Вблизи же пропускной способности, при $R \simeq C$, $p = \frac{d}{n} = \frac{2t+1}{n} \simeq 2\frac{t}{n}$.

Иными словами, вблизи пропускной способности нужно исправлять порядка d ошибок, а не $t = \frac{d-1}{2}$.

7.8. Верхняя граница Хэмминга 28-3

Доказать, что мощность M, $(k = \log M)$ q-ичного $[n, k, d]_q$-кода с минимальным расстоянием $d = 2t + 1$ не превышает границы Хэмминга

$$M \leq \frac{q^n}{|V_q(n, t)|}.$$

$$R \leq 1 - h_q(\delta/2),$$

где $R = \frac{\log_q M}{n} = \frac{k}{n}$, $\delta = \frac{d}{n}$, а

$$h_q(\delta) = -\delta \log_q \delta - (1-\delta)\log_q(1-\delta) + \delta \log_q(q-1).$$

Построить график зависимости $R(\delta)$ для $q = 2$. Привести пример совершенного двоичного кода, для которого граница Хэмминга выполняется с равенством.

Решение

Пусть имеется M слов. При $d = 2t + 1$ проведенные вокруг них сферы радиуса t не перекрываются. Полный объем этих сфер не превышает q^n. Поэтому

$$MV_q(n,t) \leq q^n.$$

Асимптотическая граница вытекает из оценки $|V_q(n,t)| \sim 2^{nh(\frac{t}{n},q)} \sim 2^{nh(\delta/2,q)}$. Пример - двоичный код Хэмминга.

7.9. Верхняя граница Синглтона 26-3

Доказать, что мощность M, $(k = \log_q M)$ q-ичного $[n,k,d]_q$-кода с минимальным расстоянием $d = 2t+1$ не превышает границы Синглтона:

$$M \leq q^{(n-d+1)}.$$

Построить асимптотическую границу

$$R \leq (1-\delta),$$

где $R = \frac{\log_q M}{n}$, $\delta = \frac{d}{n}$. Привести пример МДР-кода, лежащего на границе Синглтона.

Решение Пусть имеется q-ичный код длины n, мощности M с расстоянием d. Поделим его на q подкодов длины $(n-1)$, выделив в каждый из подкодов все слова, начинающиеся с фиксированной буквы. Среди этих подкодов найдется код мощности $M' \geq \frac{M}{q}$, а его расстояние, по прежнему, составит не менее d. После $(n-d)$-й итерации получится код длины $n-(n-d) = d$ с расстоянием d мощностью $M' \geq \frac{M}{q^{n-d}}$. Но мощность такого кода заведомо не превышает q. Так что

$$\frac{M}{q^{n-d}} \leq M' \leq q$$

или

$$M \leq q^{(n-d+1)}; \quad k = \log_q M \leq (n-d+1); \quad R = \frac{k}{n} \leq 1 - \frac{d-1}{n} = 1 - \delta.$$

8. Линейные блочные коды

8.1. Линейное пространство 29-3

Покажите, что мощность (число элементов) n-мерного линейного пространства $L_n(F_q)$ над конечным полем F_q из q элементов составляет q^n. Какова мощность одномерного подпространства (прямой), подпространства размерности k, гиперплоскости (подпространства размерности $n-1$).

Решение Каждый элемент $x \in L_n(F_q)$ волне определяется блоком $(x_1, \ldots, x_n)$ коэффициентов разложения по базису. Количество таких блоков как раз составляет q^n. Мощность k-мерного подпространства составляет q^k, прямой – q, гиперплоскости q^{n-1}.

8.2. Отображения линейных пространств 27-3

Пусть $\varphi : L_n \to L_m$ – линейное отображение (морфизм) пространства $L_n(F_q)$ в $L_m(F_q)$. Рассмотрим его ядро

$$Ker(\varphi) = \{x \in L_n : \varphi(x) = 0\},$$

и образ

$$Im(\varphi) = \{y \in L_m : \exists x \in L_n \quad y = \varphi(x)\}.$$

Покажите, что $Ker(\varphi)$ и $Im(\varphi)$ – линейные подпространства в L_n и L_m. Покажите, что

$$\frac{\dim L_n}{\dim Ker(\varphi)} = \dim Im(\varphi),$$

где dim – размерность пространства.

Решение

Пусть $\varphi(x_1) = 0$ и $\varphi(x_2) = 0$. Тогда

$$\varphi(\alpha x_1 + \beta x_2) = \alpha\varphi(x_1) + \beta\varphi(x_2) = 0.$$

Так что $Ker(\varphi)$ – подпространство. С другой стороны, пусть $y_1 = \varphi(x_1)$ и $y_2 = \varphi(x_2)$. Тогда

$$\alpha y_1 + \beta y_2 = \alpha\varphi(x_1) + \beta\varphi(x_2) = \varphi(\alpha x_1 + \beta x_2).$$

Так что $Im(\varphi)$ – подпространство.

Ясно, что отображение φ переводит разные смежные классы $L_n/Ker(\varphi)$ в разные элементы из L_m.

8.3. Линейные формы 30-3

Пусть $\varphi : L_n \to F_q$ – линейное отображение пространства $L_n(F_q)$ в поле F_q. Покажите, что его ядро

$$Ker(\varphi) = \{x \in L_n : \varphi(x) = 0\},$$

является гиперплоскостью – подпространством размерности $n-1$. Сколько элементов в фактор-пространстве $L_n/Ker(\varphi)$, каковы их образы при отображении φ. Какие линейные отображения определяют одну и ту же гиперплоскость.

Решение

Поле F_q является одномерным линейным пространством на F_q. Так что $\dim(Im(\varphi)) = 1$. Поэтому $\dim Ker(\varphi) = n - 1$. Имеется ровно q смежных классов L_n по $Ker(\varphi)$, которые при отображении φ переходят в разные элементы поля. Ясно, что отображения φ и $\alpha\varphi$ $\alpha \in F_q$ определяют одну и ту же гиперплоскость.

8.4. Двойственное пространство

Пусть $L_n(F_q)$ – линейное пространство размерности n над полем F_q. Покажите, что множество линейных отображений $\varphi : L_n \to F_q$ образует (двойственное) линейное пространство $L_n^*(F_q)$. Какова его размерность ? Докажите, что для любого базиса $(e_1, \ldots, e_n)$ L_n можно построить двойственный базис $(f_1, \ldots, f_n)$ в L_q^*, такой что $f_j(e_k) = \delta_{j,k}$.

Решение

Любой элемент $x \in L_n$ разложим по базису: $x = \sum x_j e_j$. Координатные отображения $f_k(x) = x_k$ линейны и образуют базис двойственного пространства.

8.5. Линейные отображения и матрицы 31-3

Покажите, что любое линейное отображение $\varphi : L_k \to L_n$ линейных пространств над полем F_q можно представить (k, n) матрицей с элементами из F_q. Каков класс матриц, задающих одно и то же линейное отображение.

Решение

Зафиксируем некоторые базисы $(e_1, \ldots, e_k)$ и $(f_1, \ldots, f_n)$ в L_k и L_n. Всякое линейное отображение φ вполне определяется набором образов базисных векторов. На все прочие векторы оно продолжается по линейности.

$$\varphi(e_k) = \alpha_{k,1} f_1 + \alpha_{k,2} f_2 + \ldots \alpha_{k,n} f_n.$$

Матрица коэффициентов $\{\alpha_{k,n}\}$ однозначно характеризует φ в фиксированных базисах. Матрицы, отличающиеся умножением на невырожденную квадратную матрицу слева и справа, задают представление того же линейного отображения, в других базисах.

8.6. Линейные отображения и матрицы 32-3

Покажите, что любое линейное отображение $\varphi : L_k \to L_n$ линейных пространств над полем F_q можно представить (k, n)-матрицей с элементами из F_q. Каким свойством должна обладать эта матрица, чтобы отображение φ было наложением (отображением на), таким что $\dim Im(\varphi) = m$.

Решение

Пространство $Im(\varphi)$ – это линейная оболочка столбцов матрицы отображения – образов базисных векторов. $dim Im(\varphi) = m$, если эти столбцы линейно независимы. Достаточно существование хотя бы одной невырожденной квадратной подматрицы в прямоугольной матрице отображения.

8.7. Эквивалентные коды, порождающая матрица 33-3

Линейные коды назовем эквивалентными, если один получается из другого перестановкой слов, перестановкой координат слов и покоординатным умножением всех слов на фиксированный блок $(s_1, s_2, \ldots, s_n)$ с ненулевыми координатами: $(c_1, c_2, \ldots, c_n) \to (s_1 c_1, s_2 c_2, \ldots, s_n c_n)$. Покажите, что эквивалентные коды обладают одинаковыми $[n, k, d]_q$ параметрами. Какие преобразования порождающей матрицы дают эквивалентные коды. Покажите, что среди эквивалентных кодов всегда существует код с порождающей матрицей в систематической форме.

Решение (n, k) (n строк, k столбцов) порождающую матрицу G можно умножить справа на любую невырожденную (k, k) - матрицу (перестановка слов) и умножить слева на (n, n) перестановочную матрицу и (n, n) диагональную матрицу с блоком $(s_1, s_2, \ldots, s_n))$ на главной диагонали. Среди квадратных (k, k) подматриц проверочной матрицы заведомо существует невырожденная. Умножение справа на обратную матрицу приводит ее к единичной. Остается передвинуть эту единичную матрицу в начало перестановками строк.

8.8. Эквивалентные коды, проверочная матрица 34-3

Линейные коды назовем эквивалентными, если один получается из другого перестановкой слов, перестановкой координат слов и покоординатным умножением всех слов на фиксированный блок $(s_1, s_2, \ldots, s_n)$ с ненулевыми координатами: $(c_1, c_2, \ldots, c_n) \to (s_1 c_1, s_2 c_2, \ldots, s_n c_n)$. Покажите, что эквивалентные коды обладают одинаковыми $[n,k,d]_q$ параметрами. Какие преобразования проверочной матрицы дают эквивалентные коды. Покажите, что среди эквивалентных кодов всегда существует код с проверочной матрицей в систематической форме.

Решение $(n-k,n)$ ($n-k$ строк, n столбцов) порождающую матрицу H можно умножить слева на любую невырожденную $(n-k,n-k)$ - матрицу (перестановка слов) и умножить справа на (n,n) перестановочную матрицу и (n,n) диагональную матрицу с блоком $(s_1,s_2,\ldots,s_n)$) на главной диагонали. Среди квадратных $(n-k,n-k)$ подматриц проверочной матрицы заведомо существует невырожденная. Умножение слева на обратную матрицу приводит ее к единичной. Остается передвинуть эту единичную матрицу в начало перестановками столбцов.

8.9. Порождающая-проверочная матрицы 35-3

Пусть задана (n,k) порождающая матрица G в систематической форме. Предложить алгоритм построения систематической проверочной H матрицы этого кода.

Решение Выделим в (n,k) порождающей матрице G (k,k) единичную матрицу (сверху) и $(n-k,k)$ матрицу A. Аналогично, в $(n-k,n)$ проверочной матрице H выделим $(n-k,n-k)$ единичную матрицу (права) и $(n-k,k)$ матрицу B. Нужное тождество $HG=0$ получается при $A=-B$.

8.10. Линейные коды на границе Синглтона - проверочная матрица 36-3

Покажите, что если параметры линейного $[n,k,d]_q$-кода лежат на границе Синглтона (МДР-код с $d=n-k+1$), то все $(n-k,n-k)$ квадратные подматрицы его $(n-k,n)$ проверочной матрицы невырождены. Предложите эффективный алгоритм исправления $d-1=n-k$ стираний МДР-кодом.

Решение Пусть некая $(n-k,n-k)$ матрица вырождена. Ее столбцы линейно зависимы. Значит в коде имеется слово веса $d<n-k<n-k+1$. Это не МДР код. Неизвестные значения в стертых позициях находим, обратив выделенную этими позициями невырожденную $(n-k,n-k)$ подматрицу.

8.11. Линейные коды на границе Синглтона - проверочная матрица 37-3

Покажите, что если параметры линейного $[n,k,d]_q$-кода лежат на границе Синглтона (МДР-код с $d=n-k+1$), то все (k,k) квадратные подматрицы его (n,k) порождающей матрицы невырождены, то есть никакое ненулевое кодовое слово не может принимать нулевые значения в произвольным образом заданных k-позициях. В частности, безошибочный прием любых k координат вполне определяет кодовое слово в целом. Предложите алгоритм исправления $d-1=n-k$ стираний по проверочной матрице МДР-кода.

Решение

Рассмотрим множество всех линейных отображений $\varphi(x) : F_q^k \to F_q$ k-мерного пространства в F_q и пусть $(P_1,\ldots,P_n)$ набор из n элементов F_q^k. Множество линейных отображений $\varphi(x)$ – это k-мерное линейное пространство с некоторым базисом $\varphi_j(x)$, $j=1...k$ Множество блоков

$$(c_1,\ldots,c_n)=(\varphi(P_1),\ldots,\varphi(P_n))$$

образует линейный код, столбцами порождающей матрицы которого являются блоки $(\varphi_j(P_1),\ldots,\varphi_j(P_n))$, $j=1...k$ – образы базисных отображений. Ядром всякого линейного отображения $\varphi(x)$ является некоторая гиперплоскость в F_q^k – подпространство размерности $(k-1)$. Вес кодового слова $(\varphi(P_1), \dots, \varphi(P_n))$ – это число точек P, не лежащих в гиперплоскости, связанной с отображением φ. Для построения кода с большим расстоянием нужно выбрать в F_q^k набор точек, из которых как можно большее число d не лежит ни в какой одной гиперплоскости.

Но всякие $(k-1)$ точек лежат в какой-нибудь гиперплоскости. Так что $d \le n-(k-1) = n-k+1$. Чтобы код лежал на границе Синглтона, нужно, чтобы никакие k-точек не лежали в одной гиперплоскости. Но это означает, что никакое кодовое слово не может обращаться в нуль в заданных k позициях.

8.12. Код повторения 38-3

Покажите, что $[n,k,d]_2$ код повторения с параметрами $[n=2t+1,1,2t+1]_2$ совершенен. Какова скорость этого кода, каково число гарантированно исправляемых ошибок? Как выглядят его порождающая матрица?

Решение

Код исправляет $t=\frac{d-1}{2}$ ошибок при асимптотически нулевой скорости $R=\frac{1}{n}$. Код совершенен, поскольку в t-сферу вокруг нулевого слова входит ровно половина всех слов:

$$V(n,t) = \sum_{k=0}^{t} \binom{2t+1}{k} = 2^{n-1}$$

8.13. Код Хэмминга 39-3

Покажите, что $[n,k,d]_2$ код Хэмминга с параметрами $[n=2^m-1,n-m,3]_2$ совершенен. Какова скорость этого кода, каково число гарантированно исправляемых ошибок? Как выглядят его порождающая матрица?

Решение

Код гарантированно исправляет одну ошибку при скорости $R=\frac{n-m}{n} \to 1$. Объем 1-сферы $V(n,1)$ составляет $(n+1)=2^m$, всего имеется 2^{n-m} слов, так что суммарный объем всех 1-сфер составляет как раз 2^n.

8.14. Выкалывание 40-3

Пусть имеется линейный $[n,k,d]_q$-код с $d \ge 2$. 1. Построить из него выколотый код с параметрами $[n-1,k,d-1]_q$. 2. Построить укороченный код с параметрами $[n-1,k-1,d]_q$.

Решение

1. Исключить (выколоть) одну позицию во всех кодовых словах.Исключить (выколоть) одну позицию во всех кодовых словах.

2. Рассмотреть множество всех слов с нулем в фиксированной позиции - позиции укорочения. Это линейный подкод с минимальным расстоянием d и размерностью не менее $k-1$.