第十章 率失真理论

- 由实际生活经验我们知道，一般人们并不要求完全无失真地恢复消息。对人的心理视觉研究表明，人们在观察图像时主要是寻找某些比较明显的目标特征，而不是定量地分析图像中每个像素的亮度，或者至少不是对每个像素都等同地分析。例如观看一段视频或观察一幅图像，人们可能会关注其主要情节，对视频或图像中的细节并不是那么注意，此时便允许视频或图像有一定程度的失真。

第十章 率失真理论

- 描述一个任意的实数需要无穷比特。
- 对连续随机变量的有限表示不可能完美。
- 失真度量：随机变量和它的表示之间的距离的度量。
- 率失真理论的基本问题：对于一个给定的信源分布与失真度量，在特定的码率下，可达到的最小期望失真是多少？
- 率失真理论是进行量化、数模转换、频带压缩和数据压缩的理论基础。

量化

- 例 $X \sim \mathcal{N}(0, \sigma^2)$
- X的表示（再生点）$\hat{X}(X)$：R比特
- 失真度量: $E(X - \hat{X}(X))^2$

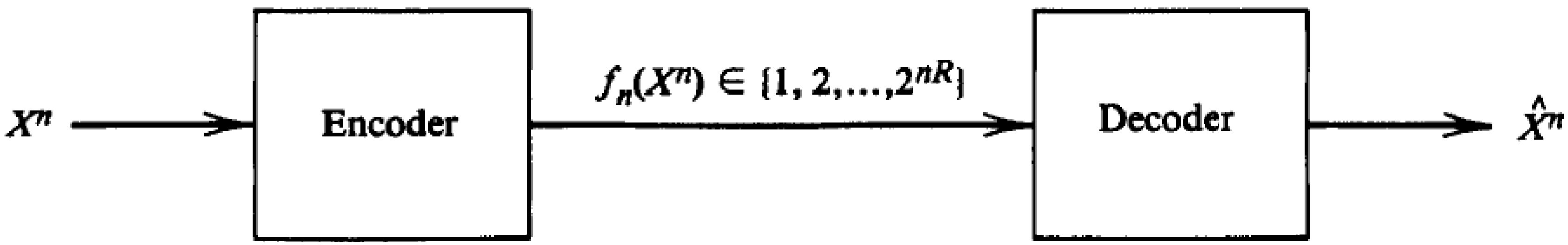
$$\hat{X}(x) = \begin{cases} \sqrt{\frac{2}{\pi}}\sigma, & \text{if } x \geq 0, \\ -\sqrt{\frac{2}{\pi}}\sigma, & \text{if } x < 0. \end{cases}$$

量化

- Lloyd算法：量化的迭代算法
  - 基于某个再生点集合，找到最优的再生区域集（在失真度量下的最邻近的区域）
  - 确定这些区域的相应最优再生点
- n个独立同分布的随机变量集合
  - 可用nR比特表示
  - 使用一个nR比特的序列来表示联合的n元随机变量，要优于使用n个R比特的序列来分别表示n个随机变量。

率失真理论模型

- 信源：$X_1, X_2, \cdots, X_n, \sim\ p(x),\quad x \in \mathcal{X}$
- 编码：$f_n(X^n) \in \{1, 2, \cdots, 2^{nR}\}$
- 译码：$\hat{X}^n \in \hat{\mathcal{X}}$



失真度量

- 失真函数(distortion function)或失真度量(distortion measure)：信源字母表与再生字母表的乘积空间到非负实数集的映射 $d:\mathcal{X}\times\hat{\mathcal{X}}\rightarrow\mathcal{R}^+$
  - ✓ 汉明（误差概率）失真：

    $$d(x,\hat{x})=\begin{cases}0, & x=\hat{x}\\ 1, & x\neq\hat{x}\end{cases}$$

    $$Ed(X,\hat{X})=\Pr(X\neq\hat{X})$$
  - ✓ 平方误差失真：

    $$d(x,\hat{x})=(x-\hat{x})^2$$
- 称失真度量是**有界的**：失真的最大值有限

$$d_{max}\triangleq\max_{x\in\mathcal{X},\hat{x}\in\hat{\mathcal{X}}}d(x,\hat{x})<\infty$$

失真度量

- 失真函数 $d(x,\hat{x})$ 是人为地规定的，给出其规定时应该考虑解决问题的需要以及失真可能引起的损失、风险和主观上感觉的差别等因素。
- $d(x,\hat{x})$是一个随机变量，它应该与信源分布函数有关，因此有必要找出在统计平均意义上信道每传送一个符号所引起失真的大小。

率失真码

- 序列间的失真：

$$d(x^n, \hat{x}^n) = \frac{1}{n}\sum_{i=1}^{n} d(x_i, \hat{x}_i)$$

- $(2^{nR}, n)$ 率失真码包括一编码函数

$$f_n : \mathcal{X} \rightarrow \{1, 2, \cdots, 2^{nR}\}$$

和一个译码（再生）函数

$$g_n : \{1, 2, \cdots, 2^{nR}\} \rightarrow \hat{\mathcal{X}}^n$$

该率失真码的失真定义为

$$D = Ed(X^n, g_n(f_n(X^n))) = \sum_{x^n} p(x^n) d(x^n, g_n(f_n(x^n)))$$

率失真函数和失真率函数

- 称率失真对(R, D)是可达的，若存在一个 $(2^{nR}, n)$ 率失真码序列 (f_n, g_n)，满足

$$\lim_{n\to\infty} Ed(X^n, g_n(f_n(X^n))) \le D$$

- 全体可达率失真对(R, D)所成的集合闭包称为信源的**率失真区域**
- 对于给定的失真D，满足(R, D)包含于信源的率失真区域中的所有码率R的下确界称为**率失真函数**（rate distortion function）R(D)
- 对于给定的码率R，满足(R, D)包含于信源的率失真区域中的所有失真D的下确界称为**失真率函数**（distortion rate function）D(R)

信息率失真函数

- 信息率失真函数：

$$R^{(I)}(D) = \min_{p(\hat{x}|x):\sum_{x,\hat{x}} p(x)p(\hat{x}|x)d(x,\hat{x})\le D} I(X;\hat{X})$$

- 定理 对于独立同分布的信源X，若公共分布为p(x)且失真函数 $d(x,\hat{x})$ 有界，那么其率失真函数与对应的信息率失真函数相等。

$$R(D) = R^{(I)}(D)$$

信息率失真函数

- 信息率失真函数：

$$R^{(I)}(D)=\min_{p(\hat{x}|x):\sum_{x,\hat{x}}p(x)p(\hat{x}|x)d(x,\hat{x})\le D} I(X;\hat{X})$$

$$\boxed{C=\max_{p(x)} I(X;Y)}$$

- 定理 对于独立同分布的信源X，若公共分布为p(x)且失真函数 $d(x,\hat{x})$ 有界，那么其率失真函数与对应的信息率失真函数相等。

$$R(D)=R^{(I)}(D)$$

信息率失真函数的理论基础

- 对于给定信源P(X)，互信息I(X;Y)是条件概率P(Y|X)的凸函数
  - 也就是说，对于一个概率分布为P(X)的信源，总可以找到某一个转移概率分布的虚拟信道，使平均交互信息量达到最小值。
  - 这个最小值是P(X)的函数，体现了信源本身的特性
  - 这个性质是信息率失真函数的理论基础

信息率失真函数的物理意义

- 信息率失真函数是在给定失真的前提下，信宿必须获得的平均信息量的最小值，是信源必须输出的最小信息率。
- 信息传输速率本质上是描述信源特性的，因此R(D)也应该是仅仅用于描述信源。
- 若信源消息经无失真编码后的信息传输速率为R，则在保真度准则下信源编码输出的信息率就是R(D)，且R(D) ＜ R
- 在保真度准则条件下的信源编码比无失真情况得到了压缩，同时R(D)是保真度条件下对信源进行压缩的极限值，亦即信源信息率可压缩的最低限度，它仅取决于信源特性和保真度要求，与信道特性无关。

信息率失真函数与信道容量的比较

- 信道容量定义为 $C = \max\limits_{p(x)} I(X;Y)$。它表示信道的最大传输能力，反映的是信道本身的特性，应该与信源无关。但平均互信息量与信源的特性有关，为排除信源特性对信道容量的影响，在所有的信源中以那个能够使平均互信息量达到最大的信源为参考，从而使信道容量仅与信道特性有关，信道不同，C亦不同。
- 信息率失真函数 $R^{(I)}(D) = \min\limits_{p(\hat{x}|x):\sum_{x,\hat{x}} p(x)p(\hat{x}|x)d(x,\hat{x})\le D} I(X;\hat{X})$

它与信道无关。为此在所有信道中以能够使平均互信息量达到最小的信道为参考，从而使信息率失真函数仅与信源特性有关，信源不同，R(D)亦不同。

引入率失真函数与信道容量的目的

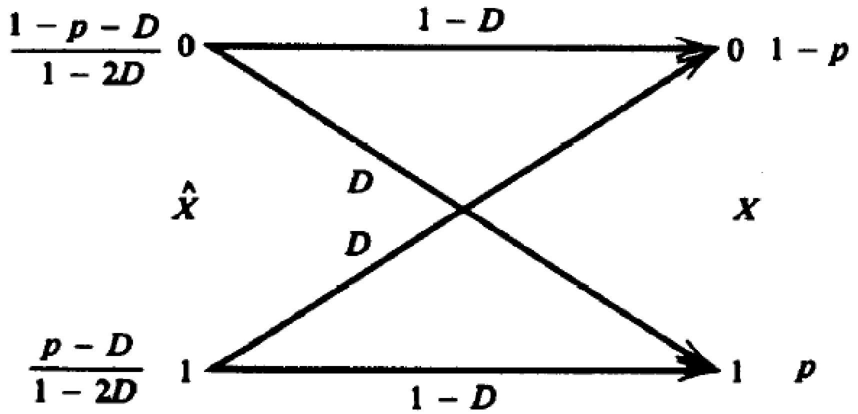
- 引入C, 是为了解决在所用信道中传送的最大信息量到底有多大的问题，它给出了信道可能传输的最大信息量，是无差错传输的上限。引入C能够为信道编码服务，或者说为提高通信的可靠性服务。

- 引入R(D)，是为了解决在允许失真度D条件下，信源编码到底能压缩到什么程度的问题，它给出了保真度条件下信源信息率可被压缩的最低限度，引入它能够为信源的压缩编码服务，或者说是为提高通信的有效性服务。

二元信源的率失真函数

- **定理** *Bernoulli(p)*信源在汉明失真度量下的率失真函数为

$$R(D)=\begin{cases}H(p)-H(D), & 0\le D\le \min\{p,1-p\},\\ 0, & D>\min\{p,1-p\}.\end{cases}$$

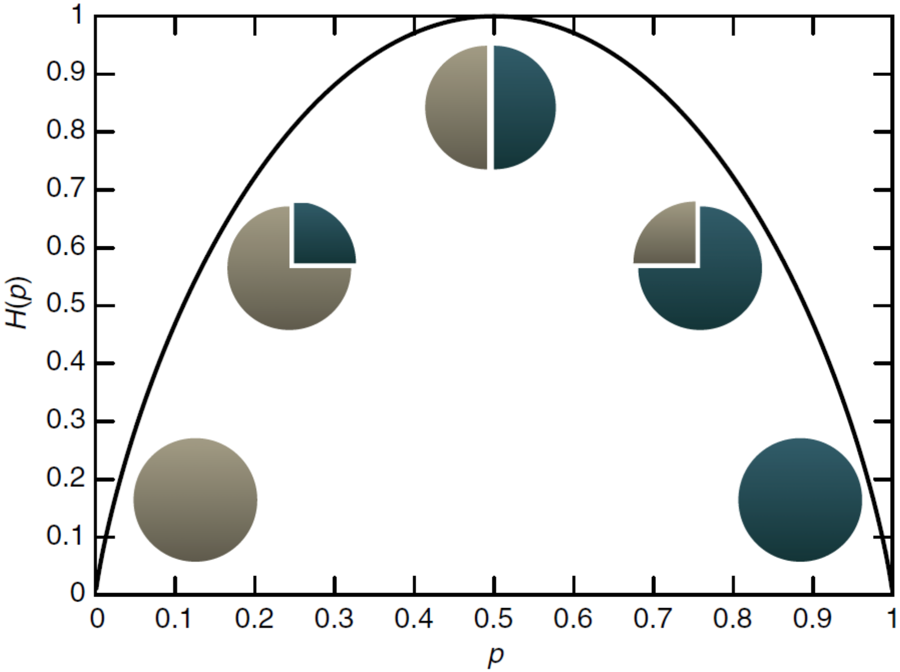


Bernoulli分布的熵

$$H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x)$$

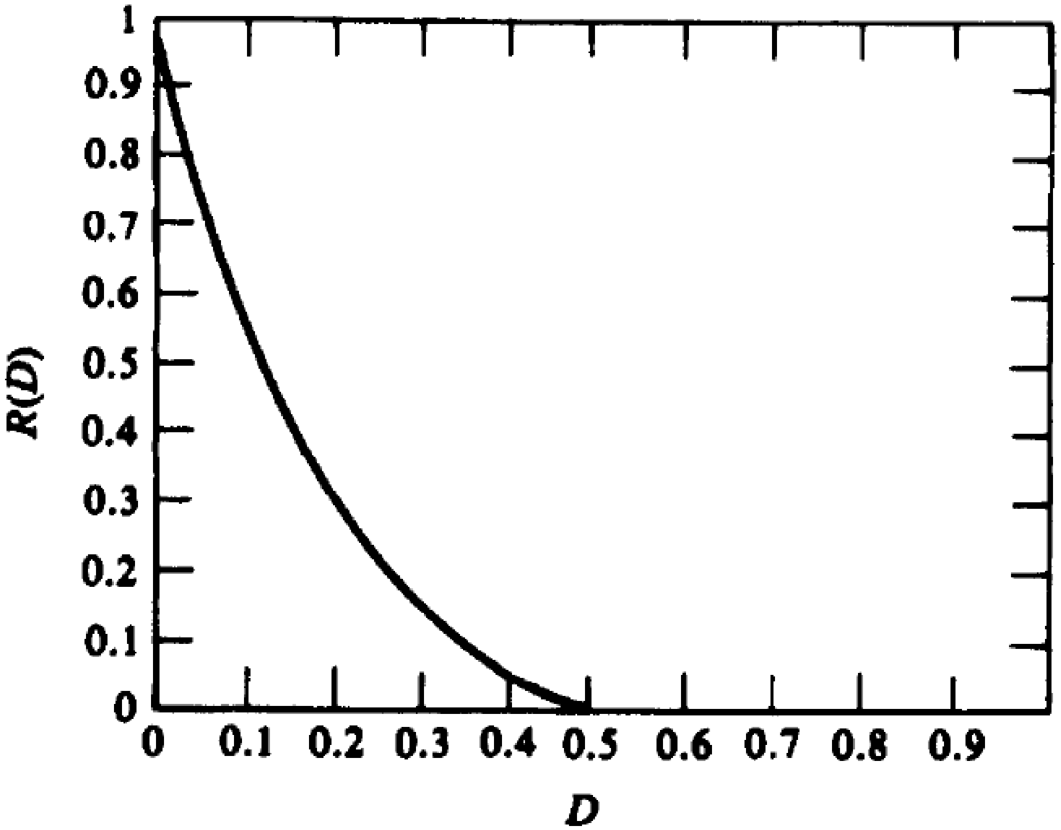
例

$$X = \begin{cases} 1 & \text{概率为}p \\ 0 & \text{概率为}1-p \end{cases}$$



$$H(X) = -p\log p - (1-p)\log(1-p) \triangleq H(p)$$

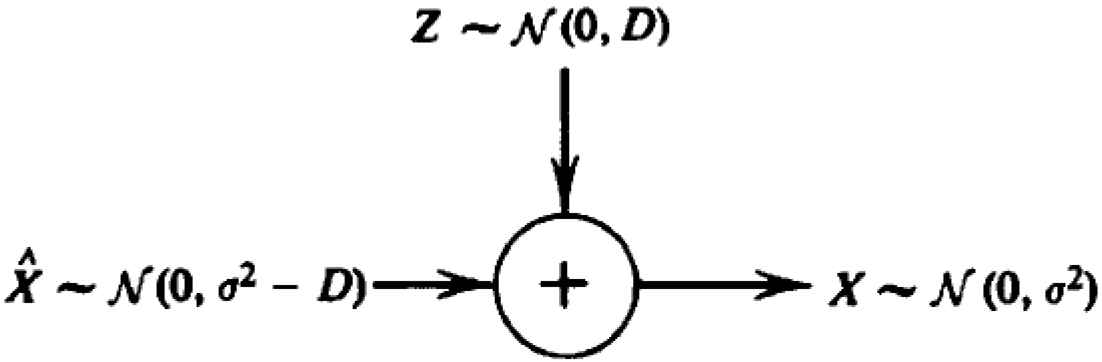
二元信源的率失真函数



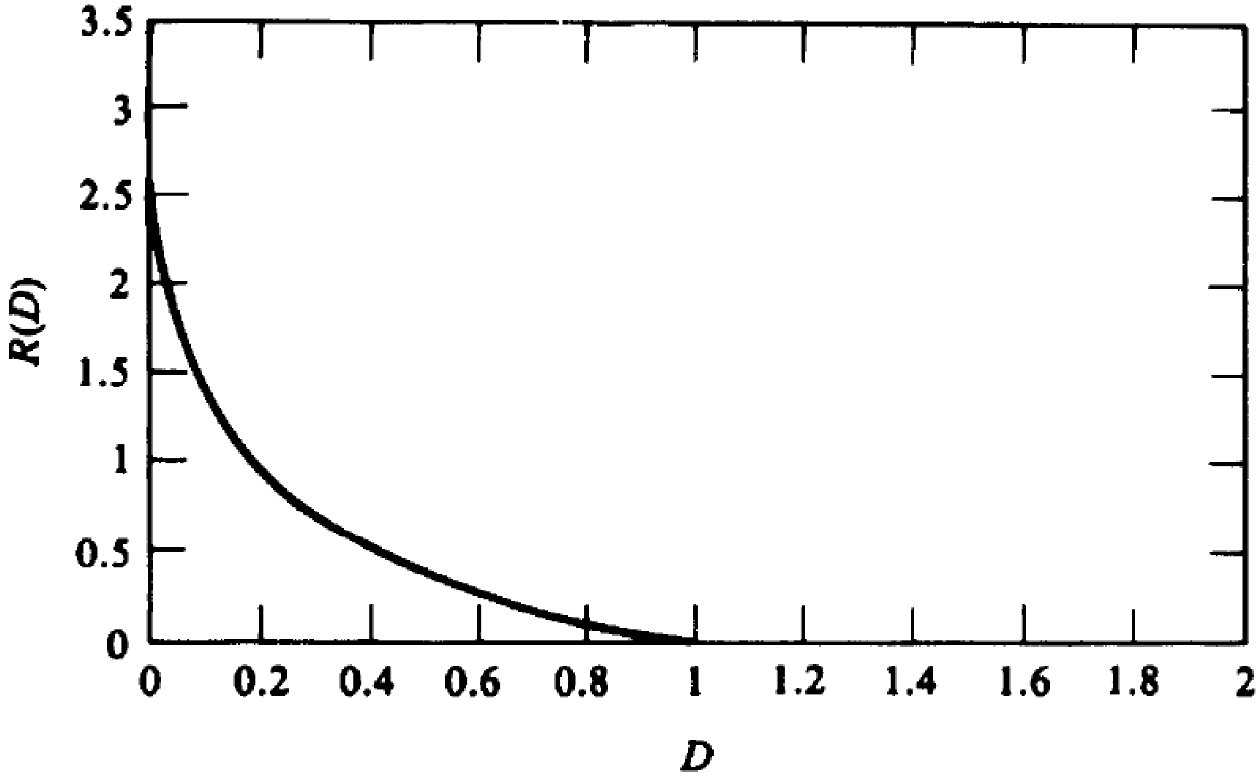
高斯信源的率失真函数

➢ 定理 一个$\mathcal{N}(0,\sigma^2)$信源在平方误差失真度量下的率失真函数为

$$R(D)=\begin{cases}\dfrac{1}{2}\log\dfrac{\sigma^2}{D}, & 0\le D\le\sigma^2,\\ 0, & D>\sigma^2.\end{cases}$$



高斯信源的率失真函数



高斯信源的失真率函数

- 失真率函数：

$$D(R) = \sigma^2 2^{-2R}$$

  - ✓ 1比特时的最佳期望平方误差： $\sigma^2/4$
- 用1比特量化高斯信源 $X \sim \mathcal{N}(0, \sigma^2)$

$$\hat{X}(x) = \begin{cases} \sqrt{\frac{2}{\pi}}\sigma, & \text{if } x \geq 0, \\ -\sqrt{\frac{2}{\pi}}\sigma, & \text{if } x < 0. \end{cases}$$

期望失真为 $\frac{\pi - 2}{\pi}\sigma^2 = 0.3633\sigma^2$

并联高斯信源的率失真

- m个独立的正态随机信源 $X_i \sim \mathcal{N}(0, \sigma_i^2)$
- 信息率失真函数的定义

$$R(D) = \min_{f(\hat{x}^m|x^m):Ed(X^m,\hat{X}^m)\le D} I(X^m;\hat{X}^m)$$

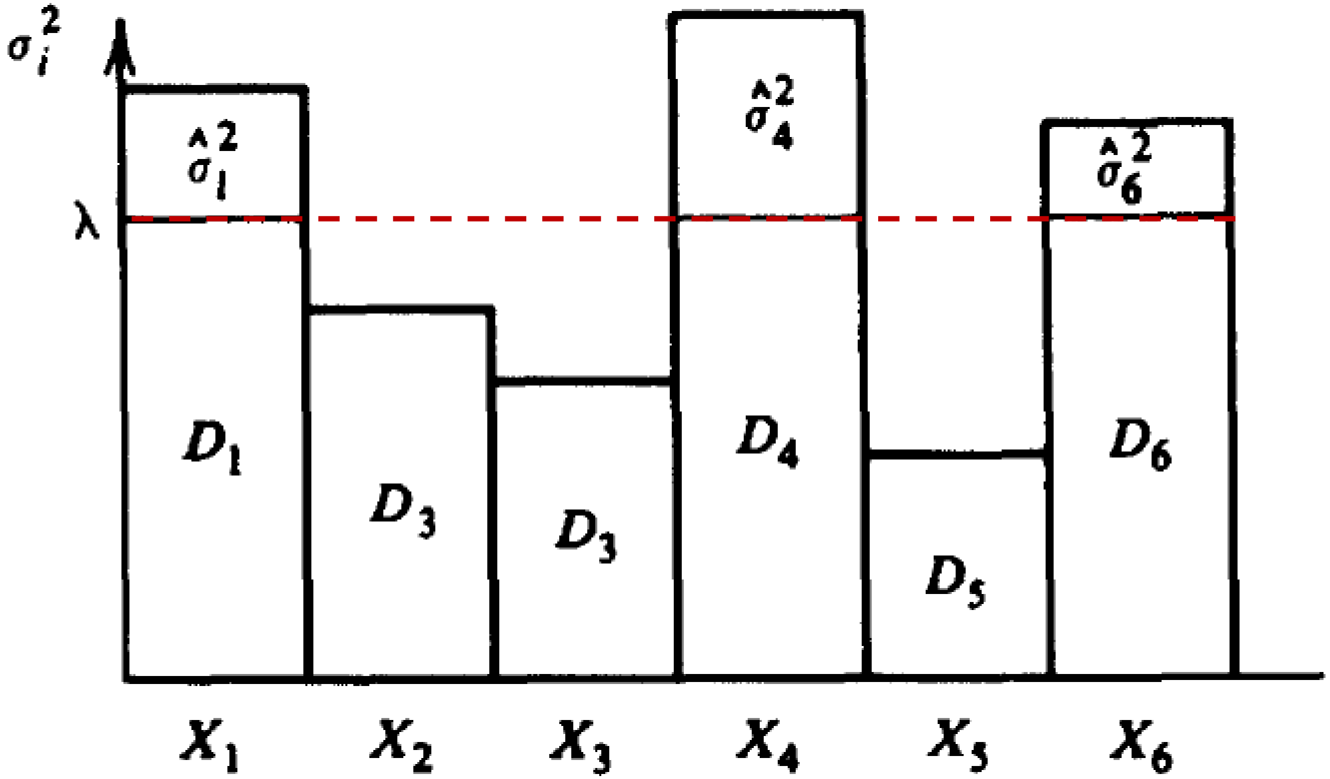
$$d(X^m,\hat{X}^m) = \sum_{i=1}^{m}(x_i - \hat{x}_i)^2$$

- 率失真函数为：

$$R(D) = \sum_{i=1}^{m} \frac{1}{2}\log\frac{\sigma_i^2}{D_i}$$

$$D_i = \begin{cases} \lambda, & \text{if } \lambda < \sigma_i^2, \\ \sigma_i^2, & \text{if } \lambda \ge \sigma_i^2, \end{cases}$$ 选取λ满足 $\sum_{i=1}^{m} D_i = D$

反注水法



σ_i^2
λ
$\hat{\sigma}_1^2$
$\hat{\sigma}_4^2$
$\hat{\sigma}_6^2$
D_1
D_3
D_3
D_4
D_5
D_6
X_1
X_2
X_3
X_4
X_5
X_6

率失真定理的逆定理

- 率失真函数R(D)是关于D的非增凸函数。

$$R(D) = \min_{p(\hat{x}|x):\sum_{x,\hat{x}} p(x)p(\hat{x}|x)d(x,\hat{x})\le D} I(X;\hat{X})$$

- 定理 率失真定理的逆定理：对于失真度量 $d(x,\hat{x})$，且i.i.d.服从p(x)的任何信源X，以及失真小于D的任何一个 $(2^{nR},n)$ 率失真码，码率必定满足

$$R \ge R(D)$$

带失真的信源信道分离定理

➢ 定理 令 $V_1, V_2, \cdots, V_n$ 为有限个独立同分布字母表的信源，编码为容量C的离散无记忆信道中的n个输入字符序列 X^n 。而信道的输出 Y^n 映射为重构字母表 $\hat{V}^n = g(Y^n)$ 。令

$$D = Ed(V^n, \hat{V}^n) = \frac{1}{n}\sum_{i=1}^{n} Ed(V_i, \hat{V}_i)$$

为由该组合信源与信道编码方案构成的平均失真。该失真D可达当且仅当C>R(D)成立。

失真典型序列

- 失真典型序列：

$$\left|-\frac{1}{n}\log p(x^n)-H(X)\right|<\epsilon$$

$$\left|-\frac{1}{n}\log p(\hat{x}^n)-H(\hat{X})\right|<\epsilon$$

$$\left|-\frac{1}{n}\log p(x^n,\hat{x}^n)-H(X,\hat{X})\right|<\epsilon$$

$$\left|d(x^n,\hat{x}^n)-Ed(X,\hat{X})\right|<\epsilon$$

- 失真典型集：$A_{d,\epsilon}^{(n)} \subset A_{\epsilon}^{(n)}$

失真典型序列的性质

- 设独立同分布的序列 $(X_i, \hat{X}_i) \sim p(x, \hat{x})$，则当 $n \to \infty$ 时，

$$\Pr(A_{d,\epsilon}^{(n)}) \to 1$$

- 对任意 $(x^n, \hat{x}^n) \in A_{d,\epsilon}^{(n)}$，

$$p(\hat{x}^n) \geq p(\hat{x}^n|x^n) 2^{-n(I(X;\hat{X})+3\epsilon)}$$

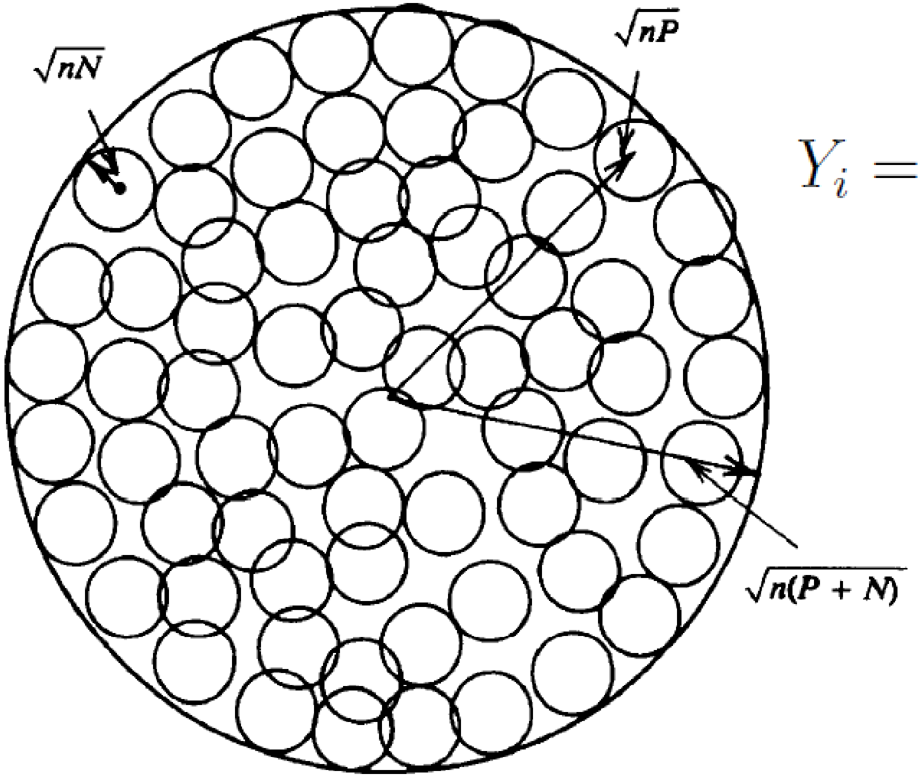
- 对 $0 \leq x, y \leq 1, n > 0$ ，

$$(1 - xy)^n \leq 1 - x + e^{-yn}$$

率失真函数的可达性

- 定理 率失真函数的可达性（香农第三定理，有失真信源编码定理）：对于任意的D和任意的R>R(D)，具有码率R和失真D的率失真码序列存在。
  - ✓ 码簿的生成
  - ✓ 编码
  - ✓ 译码
  - ✓ 失真计算
  - ✓ 误差概率计算

高斯信道信道编码定理的证明



$$Y_i = X_i + Z_i, \quad Z_i \sim \mathcal{N}(0, N)$$

率失真函数的可达性

- 香农第三定理仍然只是个存在性定理，至于最佳编码方法如何寻找，定理中并没有给出，因此有关理论的实际应用有待于进一步研究。
  - 如何计算符合实际信源的信息率失真函数R(D)?
  - 如何寻找最佳编码方法才能达到信息压缩的极限值R(D)?
  - 这是该定理在实际应用中存在的两大问题，它们的彻底解决还有赖于继续的努力。
  - 尽管如此，香农第三定理毕竟对最佳限失真信源编码方法的存在给出了肯定的回答，它为今后人们在该领域的不断深入探索提供了坚定的信心。

信息的价值

- 信息的客观评价：香农信息论的信息量；
- 信息的主观评价：重要性因人而异；
- 信息的价值评价：把平均失真理解为平均损失，信息的获得->平均失真的减少->平均损失的减少->信息价值的体现

例 某工厂生产：合格品—x_1，$p(x_1)=0.99$

废　品—x_2，$p(x_2)=0.01$

检验：合格品—y_1，合格品报废—损失1元

废　品—y_2，废品出厂—损失100元

信息的价值

例 模型：$\begin{bmatrix} \boldsymbol{X} \\ p(x_i) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ 0.99 & 0.01 \end{bmatrix}$， d $= \begin{bmatrix} 0 & 1 \\ 100 & 0 \end{bmatrix}$

1. 产品未经检验全部出厂：损失1元；
2. 产品未经检验全部报废：损失0.99元；
3. 检验完全正确：需要H(0.99)=0.081bit信息量，避免了0.99元的损失，信息价值12.2元/bit；
4. 检验有一定误差(设错判概率为0.1)：需要0.025bit信息量，避免了0.791元的损失，信息价值31.6元/bit

I(X;Y)=H(Y)-H(Y|X)=H(0.892)-H(0.9)=0.025

D=0.99*0.1*1+0.01*0.1*100=0.199