1. $I(X;\widehat{X}) = H(X) - H(X|\widehat{X}) = 1 - H(X|\widehat{X})$。因此 $\min I(X;\widehat{X})$ 等价于极大化 $H(X|\widehat{X})$。对于有限的 D，满足 $\mathbb{E}[d(X,\widehat{X})] \le D$ 必有 $\Pr(X=0,\widehat{X}=1)=0$ 因此 $\Pr(\widehat{X}=1|X=0)=0$。此时 $\mathbb{E}[d(X,\widehat{X})] \le D \Rightarrow \Pr(X=1,\widehat{X}=0) \le D \Rightarrow \Pr(\widehat{X}=0|X=1) \le 2D$

$$\begin{aligned} H(X|\widehat{X}) &= -\sum_{x,\hat{x}\in\{0,1\}} p(x,\hat{x}) \log \frac{p(x,\hat{x})}{p(\hat{x})} \\ &= -\sum_{\hat{x}=0,x\in\{0,1\}} p(x,\hat{x}) \log \frac{p(x,\hat{x})}{p(\hat{x})} \\ &= P(\widehat{X}=0)H(X|\widehat{X}=0) \end{aligned}$$

设 $q=\Pr(\widehat{X}=0|X=1)$，则 $H(X|\widehat{X}) = \frac{1+q}{2}h(\frac{1}{1+q})$，其中 h 是二元熵函数。可以证明 $H(X|\widehat{X})$ 是关于 q 的增函数，因此当 $D \le \frac{1}{2}$ 时取 $q=2D$ $H(X|\widehat{X})$ 达到最大。当 $D \ge \frac{1}{2}$ 时，取 $q=1$，此时 $H(X|\widehat{X})=1$。因此

$$R(D) = \begin{cases} 1-\frac{1+2D}{2}h(\frac{1}{1+2D}) & D \le \frac{1}{2} \\ 0 & D > \frac{1}{2} \end{cases}$$

2.

$$\begin{aligned} I(X;\widehat{X}) &= h(X) - h(X|\widehat{X}) \\ &= h(X) - h(X-\widehat{X}|\widehat{X}) \\ &\ge h(X) - h(X-\widehat{X}) \end{aligned}$$

因为 $\mathbb{E}[(X-\widehat{X})^2] = \mathbb{E}[d(X,\widehat{X})] \le D$，由最大熵分布可得 $h(X-\widehat{X}) \le \frac{1}{2}\log 2\pi eD \Rightarrow I(X;\widehat{X}) \ge h(X) - \frac{1}{2}\log 2\pi eD$。对不等式左边关于转移概率取最小值即得 $h(X) - \frac{1}{2}\log 2\pi eD \le R(D)$。

对于上界，考虑 $\widehat{X} = \frac{\sigma^2-D}{\sigma^2}(X+Z)$，其中 $Z \sim N(0,\frac{D\sigma^2}{\sigma^2-D})$，且与 X 相互独立。则

$$\begin{aligned} I(X;\widehat{X}) &= h(\widehat{X}) - h(\widehat{X}|X) \\ &= h(\widehat{X}) - h(\frac{\sigma^2-D}{\sigma^2}Z) \\ &= h(\widehat{X}) - \frac{1}{2}\log 2\pi e\frac{(\sigma^2-D)D}{\sigma^2} \end{aligned}$$

因为 $\mathbb{E}[\widehat{X}]=0, \mathbb{E}[\widehat{X}^2]=\sigma^2-D$ 所以 $h(\widehat{X}) \le \frac{1}{2}\log 2\pi e(\sigma^2-D) \Rightarrow R(D) \le I(X;\widehat{X}) \le \frac{1}{2}\log\frac{\sigma^2}{D}$

因为在相同方差条件下高斯信源 $R(D)$ 最大，相同失真度下需要更多比特编码，因此更难描述。

3. 由 Shannon 率失真下界 $R(D) \geq H(X) - \phi(D)$ 。对于 X 是均匀分布且失真矩阵各行各列互为排列组合，可取到下界。这里 $\phi(D) = \max_{\boldsymbol{p}} H(\boldsymbol{p})$ 满足约束 $\sum_{i=1}^{m} p_i = D$ 且 $\sum_{i=1}^{2m} p_i = 1$。当 $D \geq \frac{1}{2}$ 时，取分布 $p_i = \frac{1}{2m}$ 此时 $\phi(D) = H(X) \Rightarrow R(D) = 0$；当 $D < \frac{1}{2}$ 时，取分布 $p_i^* = \begin{cases} \frac{D}{m}, & i = 1, \ldots, m \\ \frac{1-D}{m}, & i = m+1, \ldots, 2m \end{cases}$ 。由 $D(\boldsymbol{p} \| \boldsymbol{p}^*) \geq 0 \Rightarrow H(X) \leq -D \log \frac{D}{m} - (1-D) \log \frac{1-D}{m}$ 。因此

$$R(D) = \begin{cases} \log(2m) + D \log \frac{D}{m} + (1-D) \log \frac{1-D}{m}, & D < \frac{1}{2} \\ 0, & D \geq \frac{1}{2} \end{cases}$$

4. 降低了 $R(D)$，因为参数空间变大了。

5. 根据最大熵原理，$f(x) = \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln x] = x^{-\lambda_2} \exp[-\lambda_0 - \lambda_1 x]$ 其中参数 $\lambda_i, i = 0, 1, 2$ 根据约束条件 $\int x f(x) dx = \alpha_1, \int (\ln x) f(x) dx = \alpha_2, \int f(x) = 1$ 确定。