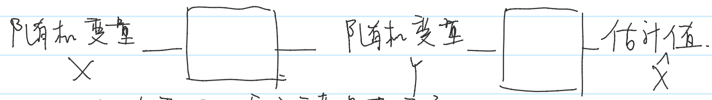
信息论与数据估计

2020年4月3日 0:09

1. 数学模型的建立

通信中，我们进行数据估计通常是在信道中。一般是编码，传输，译码。那么由于无线信道上噪声一般会引起误码。通过框图，一般则是。



那么我们通过这个模型可以定义错误概率

$$P_e = Pr\{\hat{X} \neq X\}$$

2. Fano不等式

我们希望通过信息论对于错误进行定界，对此我们有一个非常经典的不等式，Fano不等式

定理如下：

对于任意估计 $\hat{X}$，满足 $X \to Y \to \hat{X}$ 且 $P_e = Pr(X \neq \hat{X})$

$$H(P_e) + P_e \log|X| \geq H(X|\hat{X}) \geq H(X|Y)$$

这个定理的证明也是非常巧妙的。

证明如下：

① 首先我们定义一个针对二进制的差分码 $E = \begin{cases} 0, & X = \hat{X} \\ 1, & X \neq \hat{X} \end{cases}$，

那么 $H(E) = H(P_e) = P_e \log\frac{1}{P_e} + (1-P_e)\log\frac{1}{1-P_e}$

$$H(E, X|\hat{X}) = H(X|\hat{X}) + H(E|\hat{X}, X) = H(X|\hat{X})$$

$\hat{X}$ 和 X 已知，即我们已知 E，则 $H(E|\hat{X}, X) = 0$。但知道 $\hat{X}$，E 并不知道 X，因为 X 有很多个。

$$H(E, X|\hat{X}) = H(E|\hat{X}) + H(X|\hat{X}, E) = H(E|\hat{X}) + P_e H(X|\hat{X}, X \neq \hat{X}) + (1-P_e) H(X|\hat{X}, X = \hat{X})$$

$$\leq H(P_e) + P_e H(X|\hat{X}, X \neq \hat{X}) \leq H(P_e) + P_e \log|X|$$ 于是第一部分得证

② $H(X|\hat{X}) = H(X) - I(X;\hat{X})$　　　$H(X|Y) = H(X) - I(X;Y)$

要证 $H(X|\hat{X}) \geq H(X|Y)$ 即证 $I(X;Y) \geq I(X;\hat{X})$

对此我们也有一种很巧妙的证明，利用了互信息的可加性

$$I(X;Y,\hat{X}) = I(X;Y) + I(X;\hat{X}|Y) \text{ 且 } I(X;Y,\hat{X}) = I(X;\hat{X}) + I(X;Y|\hat{X})$$

那么互信息是大于等于0的，但 $I(X;\hat{X}|Y) = 0$。因为在 Y 已知的前提下，X 和 $\hat{X}$ 是两个独立的事情。

所以有 $H(X|\hat{X}) \geq H(X|Y)$，因为 $I(X;Y) \geq I(X;\hat{X})$

那么整个Fano不等式说明了这样一件事，数据越处理，互信息会减少，信息熵则会增大。而我们则对此也有一个上界即 $H(P_e) + P_e \log|X|$。