## **Sufficient Statistics**

Consider a family of probability distributions  $P_X^{\theta}$  which is parametrized by  $\theta$ . Let T(X) be any statistic (i.e. a function of sample X ). It then holds that

$$heta o X o T(X)$$
.

Hence, by the data-processing inequality, it holds that  $I(\theta;X) \geq I(\theta;T(X))$ , with equality if  $I(\theta;X\mid T(X))=0$ , or in other words, equality holds if  $\theta \leftrightarrow T(X) \leftrightarrow X$  is also a Markov chain. As we want to make sure that our statistic does not lose any information about the parameter  $\theta$ , we define the following.

## **Definition: Sufficient statistic**

 $\mathsf{T}(\mathsf{X})$  is a sufficient statistic if  $P_{X|T(X)}$  is independent of  $\theta$  for any distribution of  $\theta$ .

## **Example: Coin Flips**

Let  $X_1,X_2,\ldots,X_n$  be iid coinflips, i.e. Bernoulli(p) variables and let  $T(X_1,X_2,\ldots,X_n)=\sum_{i=1}^n X_i$  be a statistic. We have that

$$p o X_1,\ldots,X_n o \sum_{i=1}^n X_i=T(X_1,\ldots,X_n).$$

The probability of a particular outcome  $x_1 \dots x_n$  is given by

$$P_{X_1 \dots X_n}(x_1, \dots, x_n) = p^{T(x)} (1-p)^{n-T(x)}.$$

Observe that given that the number of 1's is T(x), all strings with that property are evenly likely (and therefore independent of p). Hence  $T(X) = \sum_{i=1}^{n} X_i$  is a sufficient statistic according to the definition above.

created: 2018-12-12