Definition 1 (Sufficient Statistic)**.** *Let $\{X_i\}_{i=1}^n$ be iid samples from a distribution with probability mass/density function $f(x;\theta)$ where θ is a parameter of the distribution. Let $Y = g(X_1,...,X_n)$ be a statistic. If*

$$\mathbf{P}\{X_1 = x_1,...,X_n = x_n \,|\, Y = y\} \tag{1}$$

*does not depend on θ, then Y is said to be a **sufficient statistic** for θ or **sufficient** for θ.*

Remark 1. *If a sufficient statistic for a parameter is found, then every information about such parameter that can be inferred from the data is capture in that statistic. That is, any other information (e.g. variation of the data set) does not provide any further information about the parameter.*

Example 1. *Let $\{X_i\}_{i=1}^n \overset{iid}{\sim}$ Bernoulli(p) and a statistic $Y = \sum_{i=1}^n X_i$. Consider the expression (1):*

$$
\begin{aligned}
\mathbf{P}\{X_1 = x_1,...,X_n = x_n \,|\, Y = y\} &= \frac{\prod_{i=1}^n p^{x_i}(1-p)^{1-x_i}}{\binom{n}{y}p^y(1-p)^{n-y}} \\
&= \frac{p^{\sum_{i=1}^n x_i}(1-p)^{\sum_{i=1}^n 1-x_i}}{\binom{n}{y}p^y(1-p)^{n-y}} \\
&= \frac{p^y(1-p)^{n-y}}{\binom{n}{y}p^y(1-p)^{n-y}} = \frac{1}{\binom{n}{y}}.
\end{aligned}
$$

Notice that the result does not depend on p. Therefore, Y is a sufficient statistic for p.

Theorem 1 (Fisher-Neyman Factorisation)**.** *Y is sufficient for θ iff*

$$f(x_1,...,x_n;\theta) = \phi(Y;\theta)h(x_1,...,x_n)$$

where $\phi(\cdot)$ is a function that depends on x_i only through Y and $h(\cdot)$ is a function that does not depend on θ.

References

[1] Explanation and examples of sufficient statistics. https://www.youtube.com/watch?v=G5pLCoPXr9g