Example. Suppose $X_1, \ldots, X_n$ is a random sample (i.i.d.) from Poisson(λ), $\lambda > 0$. Is $T(X) = \sum_{i=1}^n X_i$ sufficient for λ? Note that

$$\begin{aligned} f(x_1, \ldots, x_n|\lambda) &= \prod_{i=1}^n \exp(-\lambda)\frac{\lambda^{x_i}}{x_i!} = \exp(-n\lambda)\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \\ &= \left(\exp(-n\lambda)\lambda^{\sum_{i=1}^n x_i}\right)\frac{1}{\prod_{i=1}^n x_i!}. \end{aligned}$$

Take $g(t, \lambda) = \exp(-n\lambda)\lambda^t$ and $h(x_1, \ldots, x_n) = \frac{1}{\prod_{i=1}^n x_i!}$ to satisfy the factorization theorem.

Example. Let $X_1, \ldots, X_n$ be a random sample (i.i.d.) from $N(\mu, \sigma^2)$. What are jointly sufficient for μ and σ^2? We have,

$$\begin{aligned} f(x_1, \ldots, x_n|\mu, \sigma^2) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right) \\ &= (2\pi)^{-n/2}\sigma^{-n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right) \\ &= (2\pi)^{-n/2}\sigma^{-n}\exp\left(-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^n x_i^2 + n\mu^2 - 2\mu\sum_{i=1}^n x_i\right\}\right) \\ &= \sigma^{-n}\exp\left(-\frac{n\mu^2}{2\sigma^2} - \frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^n x_i\right)(2\pi)^{-n/2}. \end{aligned}$$

Note that $T(X_1, \ldots, X_n) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is sufficient since we can take $g((t_1, t_2), (\mu, \sigma^2)) = \sigma^{-n}\exp\left(-\frac{n\mu^2}{2\sigma^2} - \frac{1}{2\sigma^2}t_2 + \frac{\mu}{\sigma^2}t_1\right)$ and $h(x) = (2\pi)^{-n/2}$. Also, the map $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2) \to (\bar{X}, \sum_{i=1}^n (X_i - \bar{X})^2)$ is one-one, so $(\bar{X}, S^2)$ is another set of sufficient statistics.

Example. Let $X_1, \ldots, X_n$ be a random sample from the population with density, $f(x|\theta) = \frac{1}{2}\exp(-|x-\theta|)$, $-\infty < \theta < \infty$. Then $f(x_1, \ldots, x_n|\theta) = \frac{1}{2^n}\exp(-\sum_{i=1}^n |x_i - \theta|)$. What is sufficient for θ? Not much reduction of data is possible here, except for noting that the joint density can be written as $f(x_1, \ldots, x_n|\theta) = \frac{1}{2^n}\exp(-\sum_{i=1}^n |x_{(i)} - \theta|)$, where $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$ are the ordered observations. Therefore $T(X_1, \ldots, X_n) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ is sufficient. Order statistics provide data reduction whenever data is a random sample from a continuous distribution. Some models permit further reduction.

Interpretation of sufficiency. Observing $\mathbf{X} = (X_1, \ldots, X_n)$ is equivalent to observing $T(\mathbf{X})$ as far as information on θ is concerned. Given $T(\mathbf{X})$ (the distribution of which depends on θ), one can generate $\mathbf{X}' = (X'_1, \ldots, X'_n)$ from $P(\mathbf{X}|T)$ (which does not require θ, being uninformative). Then the probability distributions of $\mathbf{X}$ and $\mathbf{X}'$ are the same. Note that if two random quantities have the same probability distribution then they contain the same amount of information.

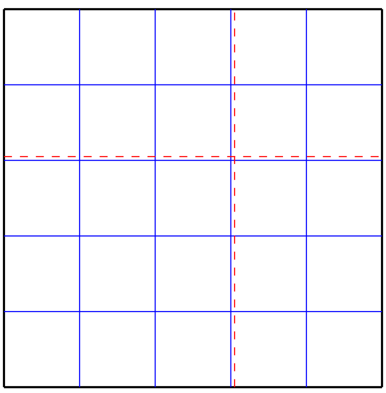
Example. $X_1, \ldots, X_n$ i.i.d Poisson(λ). Then $S = \sum_{i=1}^n X_i$ is sufficient. If $S = s$ is given from Poisson($n\lambda$), then simply generate $(X'_1, \ldots, X'_n)$ from Multinomial($s; \frac{1}{n}, \ldots, \frac{1}{n}$). It is clear that the joint distribution of $(X'_1, \ldots, X'_n)$ is the same as that of $(X_1, \ldots, X_n)$, which is i.i.d Poisson(λ).

Definition. Two statistics S_1 and S_2 are said to be equivalent if $S_1(x) = S_1(y)$ iff $S_2(x) = S_2(y)$.

Note that, if S_1 and S_2 are equivalent, then
(i) they give the same partition of the sample space,
(ii) they provide the same reduction,
(iii) they provide the same information.

Example. $S_1(X_1, \ldots, X_n) = \bar{X}$, $S_2(X_1, \ldots, X_n) = \sum_{i=1}^n X_i$. $S_1(x_1, \ldots, x_n) = S_1(y_1, \ldots, y_n)$ iff $\frac{1}{n}\sum_{i=1}^n x_i = \frac{1}{n}\sum_{i=1}^n y_i$ iff $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ iff $S_2(x_1, \ldots, x_n) = S_2(y_1, \ldots, y_n)$.

Data is a realization of the random observable $\mathbf{X}$. It is a point in the sample space. The values of $\mathbf{X}$ form the finest partition of the sample space. Any statistic $T(\mathbf{X})$ also gives a partition of the sample space. For example, $(X_1, \ldots, X_n)$ are points in $\mathcal{R}^n$. $T_1(X_1, \ldots, X_n) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ partitions $\mathcal{R}^n$ into sets where the points are permutations of each other. This is coarser than $\mathcal{R}^n$. $T_2(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^n X_i$ partitions $\mathcal{R}^n$ into sets where the points have the same average. This partition is coarser than the one provided by T_1 since permutations do not change the average.



In the figure above, the dashed, red partition is coarser than the blue partition.

Sufficient statistics gives a partition (of the sample space) which retains all the information about the parameters. Therefore, maximum possible reduction of data, or the coarsest possible partition of the sample space is desirable. How does one choose sufficient statistics?

Example. $X_1, \ldots, X_n$ i.i.d $N(0, \sigma^2)$. Then

$$f(x_1, \ldots, x_n|\lambda) = (2\pi)^{-n/2}\sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right).$$

Note from this joint density that
(i) $T_1(X_1, \ldots, X_n) = (X_1, \ldots, X_n)$ is sufficient for σ^2;
(ii) $T_2(X_1, \ldots, X_n) = (X_1^2, \ldots, X_n^2)$ is sufficient for σ^2;
(iii) $T_3(X_1, \ldots, X_n) = (X_1^2 + \cdots + X_m^2, X_{m+1}^2 + \cdots + X_n^2)$ is sufficient for σ^2;
(iv) $T_4(X_1, \ldots, X_n) = (X_1^2 + \cdots + X_n^2)$ is sufficient for σ^2.

Observe that $T_4 = g_1(T_3) = g_1(g_2(T_2)) = g_1(g_2(g_3(T_1)))$. i.e., if T is sufficient and $T = H(U)$, then U is also sufficient. Knowledge of U implies knowledge of T and hence permits reconstruction of the original data. T provides greater reduction or coarser partition unless H is one-one.

Minimal sufficiency. $T = T(X)$ is minimal sufficient if it provides the maximal amount of data reduction. i.e., for any sufficient statistics $U = U(X)$, there exists a function H such that $T = H(U)$.