A **parametric model** is a set that can be parameterized by finite number of parameters.

For a data coming from Normal distribution, we have the density

$$
\mathcal{F} = \left\{ f(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ \frac{1}{2\sigma^2}(x - \mu)^2 \right\}, \mu \in \mathbb{R}, \sigma > 0 \right\}
$$

This is a two parameter model.

In general, a parametric model is defined as,

$$
\mathcal{F} = \left\{ f(x; \theta) : \theta \in \Theta \right\}
$$
Regression, prediction, classification

- Suppose we observe pairs of data \((X_1, Y_1)\)…. \((X_i, Y_i)\)
- \(X\) is called predictor or regressor or feature or independent variable
- \(Y\) is called the outcome or the response variable or dependent variable
- The regression function “r” is defined as
  \[ r(x) = E(X \mid Y = x) \]
- The goal of predicting \(Y\) based on their \(X\) value is called prediction.
- If \(Y\) is discrete, then prediction is instead called classification
- Regression models are sometimes written as,
  \[ Y = r(X) + \varepsilon \]
Fundamental Concepts in Inference
1. Point Estimation

- Point estimation refers to providing a single best guess of some quantity of interest. The quantity of interest could be a parameter in a parametric model, CDF $F$, probability density function $f$, a regression function $r$ or prediction for future value $Y$ of some random variable.

- The point estimate of $\theta$ is denoted by $\hat{\theta}$ or $\hat{\theta}_n$

$$\hat{\theta}_n = g(X_1, \ldots, X_n)$$

$$bias(\hat{\theta}_n) = E_\theta(\hat{\theta}_n) - \theta$$
Fundamental Concepts in Inference
2. Confidence Sets

A $1 - \alpha$ confidence interval for a parameter $\theta$ is an interval $C_N = (a, b)$ where $a = a(X_1, \ldots, X_N)$ and $b = b(X_1, \ldots, X_N)$ are functions of the data such that

$$P_{\theta}(\theta \in C_N) \geq 1 - \alpha$$

for all $\theta \in \Theta$. In other words, $(a, b)$ traps $\theta$ with probability $1 - \alpha$. We call $1 - \alpha$ the coverage of the confidence interval.
3. Hypothesis Testing

In *hypothesis testing*, we start with some default theory, called a *null hypothesis*, and we ask if the data provide sufficient evidence to reject the theory. If not, we retain the null hypothesis (innocent till proven guilty). We call the opposite of the null hypothesis the *alternative hypothesis*.

A common type of hypothesis is about parameter values, where we are testing

\[ H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0 \]

or the null hypothesis versus the alternative. A hypothesis of this form is called a *two-sided test*. A *one-sided test* is about an inequality rather than an equality.
Fundamental Concepts in Inference

3. Hypothesis Testing

Let $X$ be a random variable and $\mathcal{X}$ be the range of $X$. Our test rejects the null hypothesis if some *test statistic* $T(X)$ is greater than some *critical value* $c$.

We say that we reject $H_0$ if $X \in R$, where

$$R = \{x : T(x) > c\}$$

is the *rejection region*. 