

- Y or X : a random variable
- $\mathbf{Y} = (Y_1, \dots, Y_n)$ or $\mathbf{X} = (X_1, \dots, X_n)$: a random row vector
- \mathbf{X}^T : transpose of \mathbf{X} , i.e., a random column vector.
- y observed data from Y (x is observed data from X)
- $\mathbf{y} = (y_1, \dots, y_n)$ observed data from \mathbf{Y} ($\mathbf{x} = (x_1, \dots, x_n)$ is observed data from \mathbf{X})
- $\boldsymbol{\theta} = (\theta_1, \dots, \theta_b)$: a vector contains b unknown parameters.
- iid: independent and identical distributed
- $f(y; \boldsymbol{\theta})$ (or $f(x; \boldsymbol{\theta})$): density function of Y (or X) with parameters $\boldsymbol{\theta}$
- $F(y; \boldsymbol{\theta})$ (or $F(x; \boldsymbol{\theta})$): distribution function of Y (or X) with parameters $\boldsymbol{\theta}$
- $f(\mathbf{y}; \boldsymbol{\theta})$ (or $F(\mathbf{y}; \boldsymbol{\theta})$): joint density function of \mathbf{Y} (or \mathbf{X}) with parameters $\boldsymbol{\theta}$:
- $F(\mathbf{y}; \boldsymbol{\theta})$ (or $F(\mathbf{x}; \boldsymbol{\theta})$): joint distribution function of \mathbf{Y} (or \mathbf{X}) with parameters $\boldsymbol{\theta}$
- $L(\boldsymbol{\theta}|\mathbf{y})$: the likelihood function of $\boldsymbol{\theta}$ given data \mathbf{y} .
- $\hat{\boldsymbol{\theta}}_{MLE}$: the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$.
- $\arg \max_{\boldsymbol{\theta}} g(\boldsymbol{\theta})$: The value of $\boldsymbol{\theta}$ that makes $g(\boldsymbol{\theta})$ achieve maximum.
- $\frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$: the derivative of function g with respect to $\boldsymbol{\theta}$
- \bar{y} : the sample average, i.e., $\frac{\sum_{i=1}^n y_i}{n}$
- $\prod_{i=1}^n x_i$: The cumulative product from x_1 to x_n
- g^{-1} : the inverse of $g(\cdot)$.