Probabilistic tools

Central Limit Theorem
Let $(X_n)_{n\geq 1}$ be a real and independent sequence with same law such that $\mu = \mathbb{E}[X_1]$ and $\mathbb{V}[X_1] = \sigma^2$ are defined ($\mathbb{V}[X_1] \leq +\infty$). Noting $\bar{X}_n = \frac{1}{n}(X_1 + ... + X_n)$, we have:

$$\sqrt{n}\frac{(\bar{X}_n - \mu)}{\sigma} \sim_{n\to\infty} \mathcal{N}(0,1)$$

Spectral Theorem
Let M be a symmetric matrix with real coefficients. Then it exists U orthogonal and D diagonal with real coefficients such that $M = UDU^T$.

Inferential statistics

Likelihood method

This method consists on finding the parameter that maximizes the likelihood:
$L(x_1, ..., x_n; \theta) = f(X|\theta) = \prod_{i=1}^{n} f_\theta(x_i; \theta)$ which is the product of densities across all samples.

Intuitively, we want to find the θ that maximizes a certain event, that is, obtaining some data X (which is why we have $X|\theta$).
We often use the log in order to get rid of power coefficients appearing with the product.
likelihood equation: $\frac{d}{d\theta} ln(L(x_1, ..., x_n; \theta)) = 0$

Exploratory statistics

Mahalanobis distance
Mahalanobis distance is a good alternative to Euclidean distance. For any given point x in a set X, the squared Mahalanobis distance is:

$$D^2 = (x - \mu_X)^T \Sigma^{-1} (x - \mu_X)$$

Advantage : it takes into account the data standard deviation and correlation. The more the data is dispersed, the lower the distance is. Indeed, using the inverse matrix is like if we divided the distance from the mean $(x - \mu_X)$ by the standard deviation.
Note: Euclidean distance is when $\Sigma = Id$.