📘 Probability Distributions & Naïve Bayes Summary

This document contains a comprehensive summary of important probability distributions and Naïve Bayes classifiers used in data science.

# 📊 Probability Distributions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Distribution | Type | Use Case | Mean | Variance | Real-Life Example | Parameters |
| Bernoulli | Discrete | Binary outcomes | p | p(1-p) | Coin toss | p |
| Binomial | Discrete | No. of successes in n trials | np | np(1-p) | Marketing conversions | n, p |
| Poisson | Discrete | Count of events | λ | λ | Calls per minute | λ |
| Uniform | Continuous | Equal chance outcomes | (a+b)/2 | (b-a)^2/12 | Fair die roll | a, b |
| Normal | Continuous | Natural phenomena | μ | σ^2 | Height, weight | μ, σ |
| Exponential | Continuous | Time between events | 1/λ | 1/λ² | Time between arrivals | λ |
| t-Distribution | Continuous | Small-sample inference | 0 | >1 (depends on df) | Small sample t-tests | df |
| Chi-Square | Continuous | Hypothesis testing | k | 2k | Independence test | k |
| Gamma | Continuous | Skewed waiting times | kθ | kθ² | Rainfall modeling | k, θ |
| Beta | Continuous | Probabilities [0,1] | α/(α+β) | αβ/[(α+β)^2(α+β+1)] | A/B testing | α, β |
| Multinomial | Discrete | Multiple outcomes | np₁...npk | np₁(1-p₁)... | Vote counts | n, p₁...pk |
| Multivariate Normal | Continuous | Correlated variables | μ (vector) | Σ (matrix) | Stock returns | μ, Σ |
| Dirichlet | Continuous | Proportions | αᵢ / Σα | Complex | Topic modeling | α₁, α₂, ..., αk |

# 🤖 Naïve Bayes Classifiers

Naïve Bayes is based on Bayes' Theorem and assumes feature independence. It's commonly used in classification problems.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Feature Type | Use Case | Assumption | Real-Life Example |
| Gaussian Naïve Bayes | Continuous | Iris classification | Normal distribution | Medical diagnosis |
| Multinomial Naïve Bayes | Discrete Counts | Text classification | Multinomial distribution | Spam detection |
| Bernoulli Naïve Bayes | Binary | Binary text features | Bernoulli distribution | Word presence/absence |