Probability Distributions Cheat Sheet

Alireza Miraliakbar

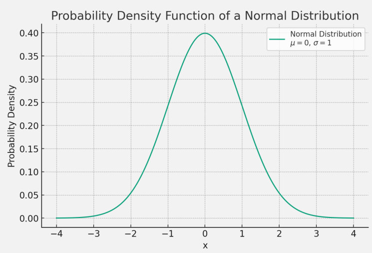
Spring 2024

Continuous Distributions

Normal (Gaussian) Distribution

Description: Described by its mean (μ) and standard deviation (σ), and has the classic "bell curve" shape.

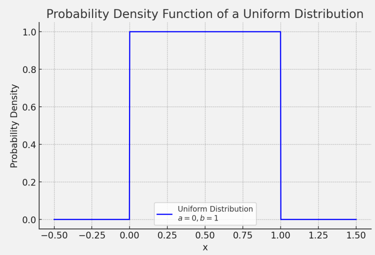
PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
Mean: μ
Variance: σ^2



Uniform Distribution

Description: Models equally likely outcomes between two boundaries, a and b. It is characterized by constant probability density over its support.

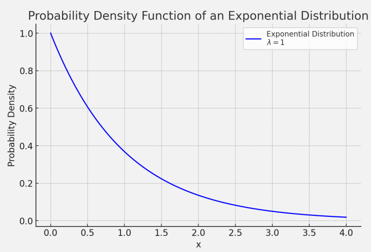
PDF: $f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$
Mean: $\mu = \frac{a+b}{2}$
Variance: $\sigma^2 = \frac{(b-a)^2}{12}$



Exponential Distribution

Description: Models the time between events in a Poisson point process, i.e., events occur continuously and independently at a constant average rate.

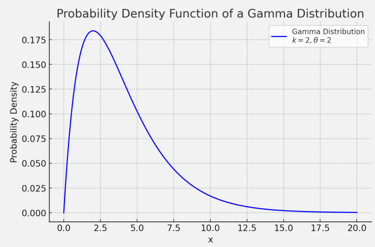
PDF: $f(x; \lambda) = \lambda e^{-\lambda x}$ for $x \geq 0$
Mean: $\frac{1}{\lambda}$
Variance: $\frac{1}{\lambda^2}$



Gamma Distribution

Description: Models the time until an event occurs a specific number of times, given a continuous and constant arrival rate. It generalizes the exponential distribution.

PDF: $f(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)}$ for $x > 0$
Mean: $\mu = k\theta$
Variance: $\sigma^2 = k\theta^2$



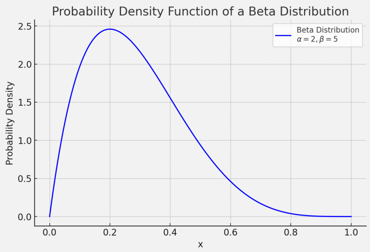
Beta Distribution

Description: Models the distribution of probabilities, representing all the possible values of probability when outcomes are distributed according to a binomial distribution. It is especially useful in Bayesian inference and is defined on the interval [0, 1].

PDF: $f(x;\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ for $0 \leq x \leq 1$

Mean: $\mu = \frac{\alpha}{\alpha+\beta}$

Variance: $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$



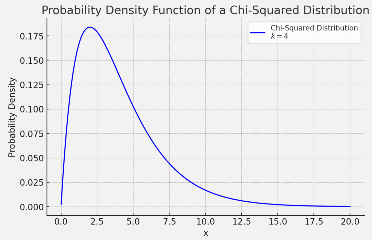
Chi-Squared Distribution

Description: A special case of the Gamma distribution, used primarily in hypothesis testing and in the construction of confidence intervals. It describes the distribution of the sum of the squares of k independent standard normal random variables.

PDF: $f(x;k) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$ for $x > 0$

Mean: $\mu = k$

Variance: $\sigma^2 = 2k$



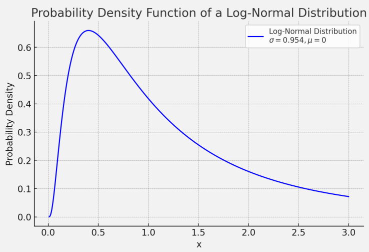
Log-Normal Distribution

Description: Models variables that are the product of many independent, identically distributed random variables. It is suitable for representing variables that cannot be negative and have a long right tail, such as income, length of time, and biological measurements.

PDF: $f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$ for $x > 0$

Mean: $\mu = e^{\mu+\frac{\sigma^2}{2}}$

Variance: $\sigma^2 = (e^{\sigma^2}-1)e^{2\mu+\sigma^2}$



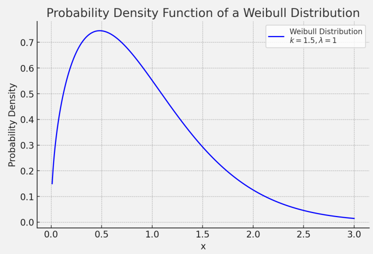
Weibull Distribution

Description: Commonly used in reliability engineering and failure analysis, the Weibull distribution is flexible, able to model various types of data. It is characterized by a scale parameter λ and a shape parameter k, which dictate the distribution's spread and form, respectively.

PDF: $f(x;k,\lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}$ for $x > 0$

Mean: $\mu = \lambda\Gamma\left(1+\frac{1}{k}\right)$

Variance: $\sigma^2 = \lambda^2\left[\Gamma\left(1+\frac{2}{k}\right) - \left(\Gamma\left(1+\frac{1}{k}\right)\right)^2\right]$



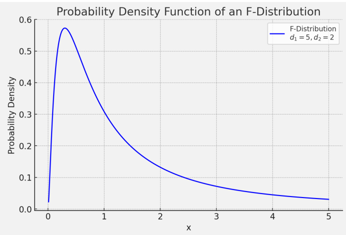
F-Distribution

Description: Used primarily in the analysis of variance (ANOVA), comparing two variances, and in regression analysis. The F-distribution arises from the division of two chi-squared distributions divided by their respective degrees of freedom.

PDF: $f(x; d_1, d_2) = \frac{\sqrt{\frac{(d_1 x)^{d_1} \cdot d_2^{d_2}}{(d_1 x + d_2)^{d_1 + d_2}}}}{x \cdot B\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$ for $x > 0$

$Mean$: $\mu = \frac{d_2}{d_2 - 2}$ for $d_2 > 2$

$Variance$: $\sigma^2 = \frac{2d_2^2(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)}$ for $d_2 > 4$



Discrete Distributions

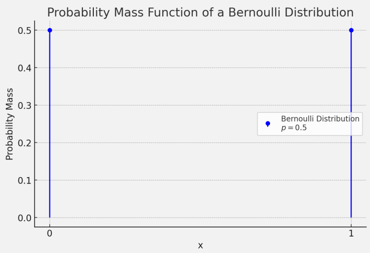
Bernoulli Distribution

Description: Models a single trial with two possible outcomes (success with probability p and failure with probability $1 - p$).

PMF: $P(X = k) = p^k(1-p)^{1-k}$ for $k \in \{0, 1\}$

Mean: $\mu = p$

Variance: $\sigma^2 = p(1-p)$



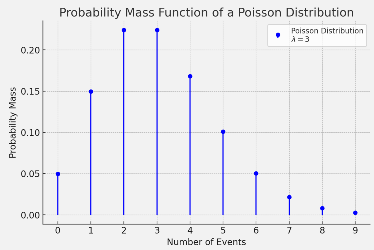
Poisson Distribution

Description: Models the number of events occurring in a fixed interval of time or space if these events happen with a known constant mean rate and independently of the time since the last event.

PMF: $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ for $k = 0, 1, 2, \ldots$

Mean: $\mu = \lambda$

Variance: $\sigma^2 = \lambda$



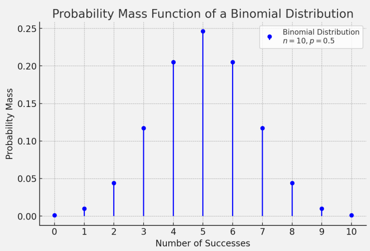
Binomial Distribution

Description: Models the number of successes in a fixed number of independent Bernoulli trials, each with the same probability of success.

PMF: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, \ldots, n$

Mean: $\mu = np$

Variance: $\sigma^2 = np(1-p)$



Geometric Distribution

Description: Models the number of trials needed to get the first success in a series of independent Bernoulli trials, each with the same probability of success p. It is a discrete distribution and one of the simplest scenarios of failure/success experiments.

PMF: $P(X = k) = (1-p)^{k-1}p$ for $k = 1, 2, \ldots$

Mean: $\mu = \frac{1}{p}$

Variance: $\sigma^2 = \frac{1-p}{p^2}$