Some example probability distributions

Discrete random variables

Continuous random variables.

① Bernoulli distribution

→ Random variable with only two possible outcomes, say success (1) or failure (0).

$$X \sim \text{Bern}(p) \; ; \; p \in [0,1]$$

↳ probability of success

$$P(X=x) = \begin{cases} p \; ; & x=1 \\ (1-p) \; ; & x=0 \\ 0 \; ; & \text{otherwise} \end{cases}$$

(2) Binomial distribution

→ generalizes the Bernoulli distribution to n independent and identical trials.

Each of the trials has 'p' as the probability of success

$$X \sim Bin(n, p)$$

represents the number of successes in n independent Bernoulli trials, each with a success probability p.

$$P(X = k) = \underbrace{{}^{n}C_{k} \, p^{k} \, (1-p)^{n-k}}_{} \quad ; \; k = 0, 1, 2, \dots, n$$

$$^{n}C_{k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

(3) Poisson distribution

The Poisson distribution models the number of events occurring in a fixed interval of time (or space), assuming each event occurs independently and at a constant average rate.

$$X \sim \text{Poisson}(\lambda)$$

λ → rate parameter : average rate of the process or event

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots$$

(4) Geometric distribution

→ models the number of Bernoulli trials needed to get the first success.

$$X \sim \text{Geom}(p) \; ; \; 0 < p \leq 1$$

$$P(X = k) = (1-p)^{k-1} p \; ; \; k = 1, 2, 3, \ldots$$

probability mass function

Continuous random variables

(1) Uniform distribution

models a continuous random variable $X \sim U(a,b)$ which has equal probability to lie anywhere in the domain $x \in (a,b)$

$$f(x) = \begin{cases} \frac{1}{b-a} \; ; & a \le x \le b \\ 0 \; ; & x < a \\ 0 \; ; & x > b. \end{cases}$$

probability density function

$$F(x) = \begin{cases} 0 \; ; & x < a \\ \frac{x-a}{b-a} \; ; & a \le x \le b \\ 1 \; ; & x > b \end{cases}$$

Cumulative distribution function

(2) Normal / Gaussian distribution

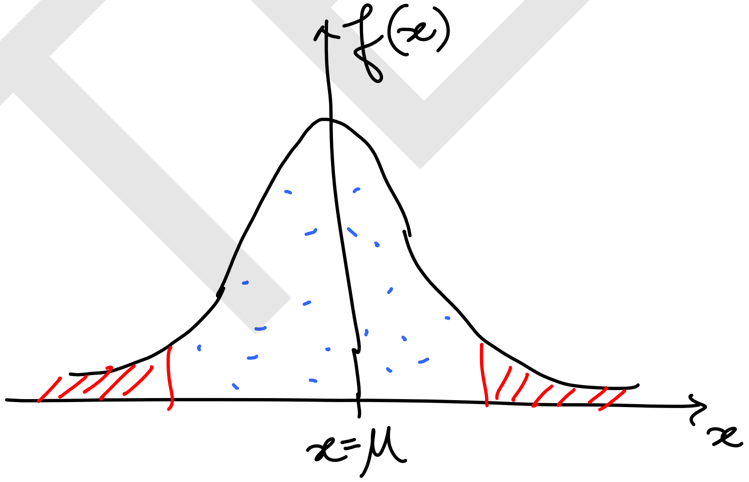
$\rightarrow \quad X \sim N(\mu, \sigma^2)$

variance

→ commonly known as a "bell curve", which represents one of the most commonly encountered distributions

→ 'Central limit theorem': the normalized sum of many independent and identical distributions tends to the normal distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



μ: mean of the distribution

σ: standard deviation

→ eg: student marks in an exam, height of people in an institution, atomic velocities in a fluid.

③ Student's t distribution

→ Determining the confidence intervals for a linear regression fit to the range in which the true mean of a parameter lies.

→ Quantifies the extent of deviation from the true mean.

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

random variable following student's t distribution

$$\bar{X} = \text{sample mean} = \boxed{\frac{1}{n} \sum_{i=1}^{n} X_i}$$

$$s^2 = \text{sample variance} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$\nu \equiv$ number of degrees of freedom $= (n-1)$

pdf of student's t distribution

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\;\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

Γ : Gamma function $\Rightarrow$ $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u}\, du = (x-1)!$

4 Exponential distribution

Represents the probability underlying the time interval between two incidences of an event.

$X \sim Exp(\lambda)$ → rate parameter

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} ; & x \geq 0 \\ 0 ; & x < 0 \end{cases}$$