Yusong Wang

Raghavendra Kommavarapu

Theo Kim

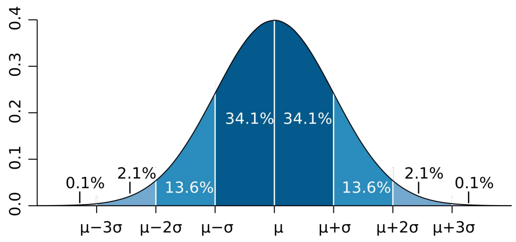
Why is the normal distribution everywhere?

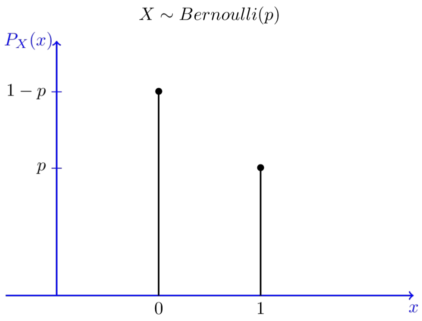
The importance of the Central Limit Theorem

Let’s talk about distributions!

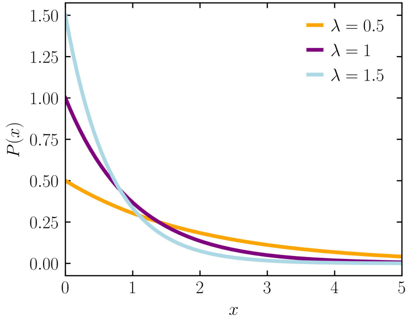
$$f(x) = \begin{cases} \lambda e^{-\lambda * x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Mathematical function that gives the probabilities of occurrence of different possible **outcomes**

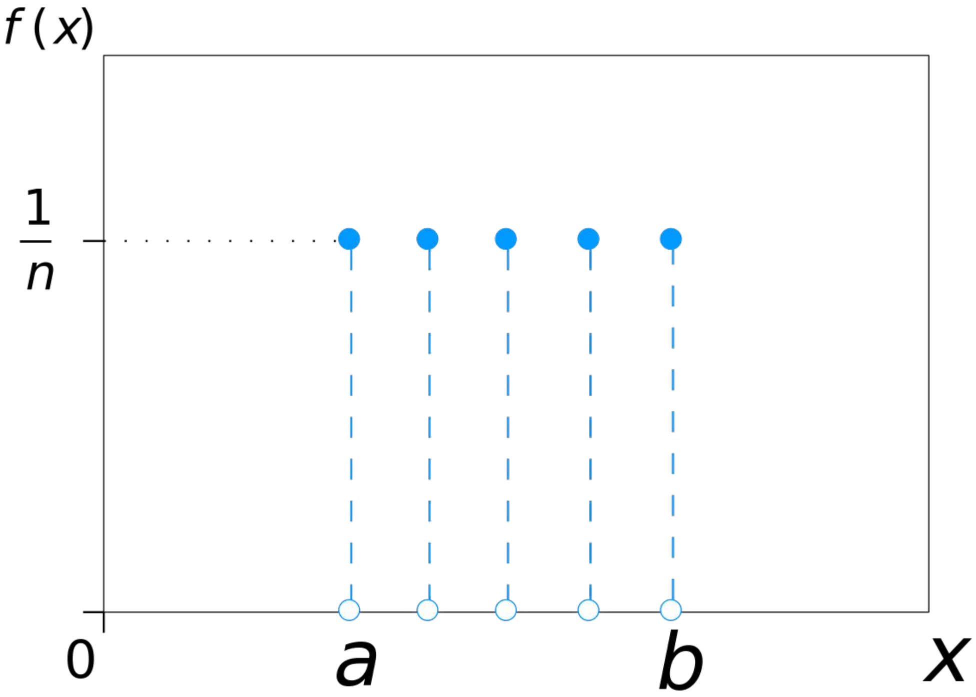




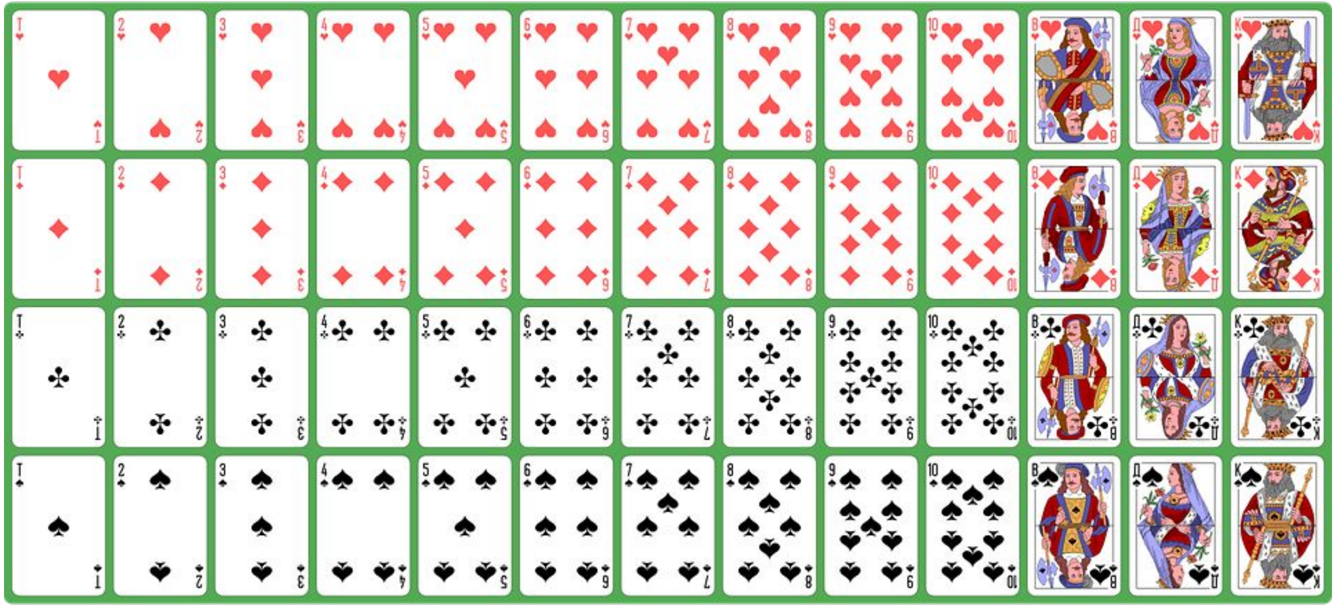
$$P(X = x) = \begin{cases} p \ \ for\ x = 1 \\ 1 - p \ \ for\ x = 0 \end{cases}$$



Uniform Distribution



Uniform Distribution

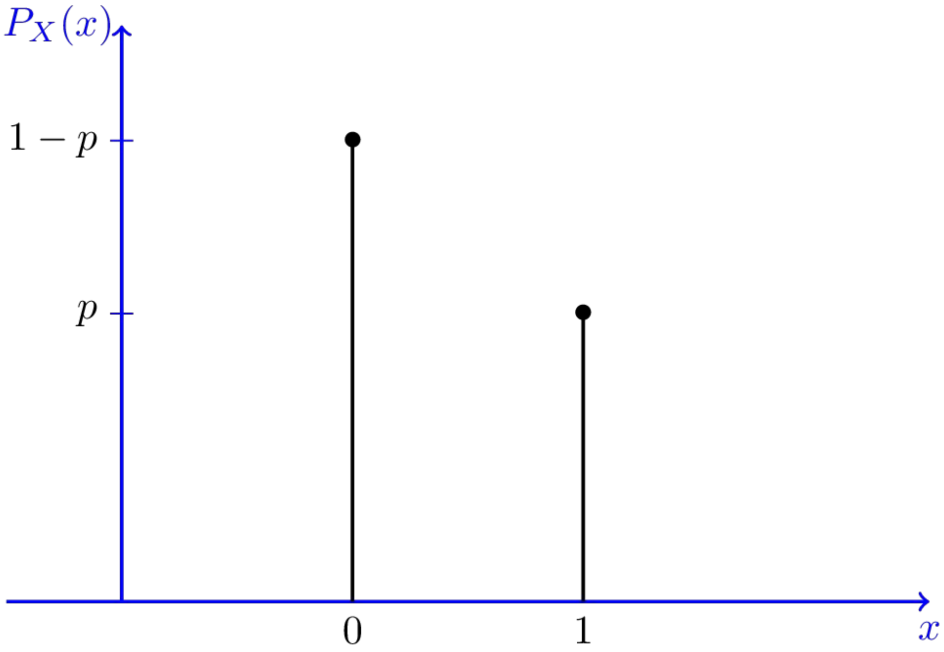


Uniform Distribution

Bernoulli Distribution

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

$X \sim Bernoulli(p)$



Bernoulli Distribution

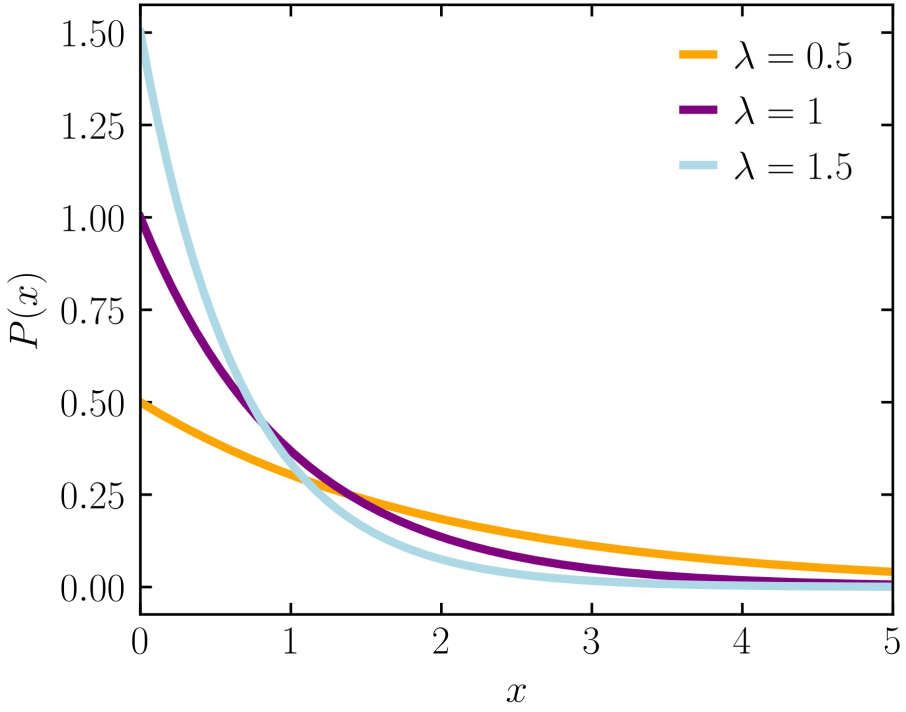
Exponential Distribution

Exponential Distribution

Exponential Distribution

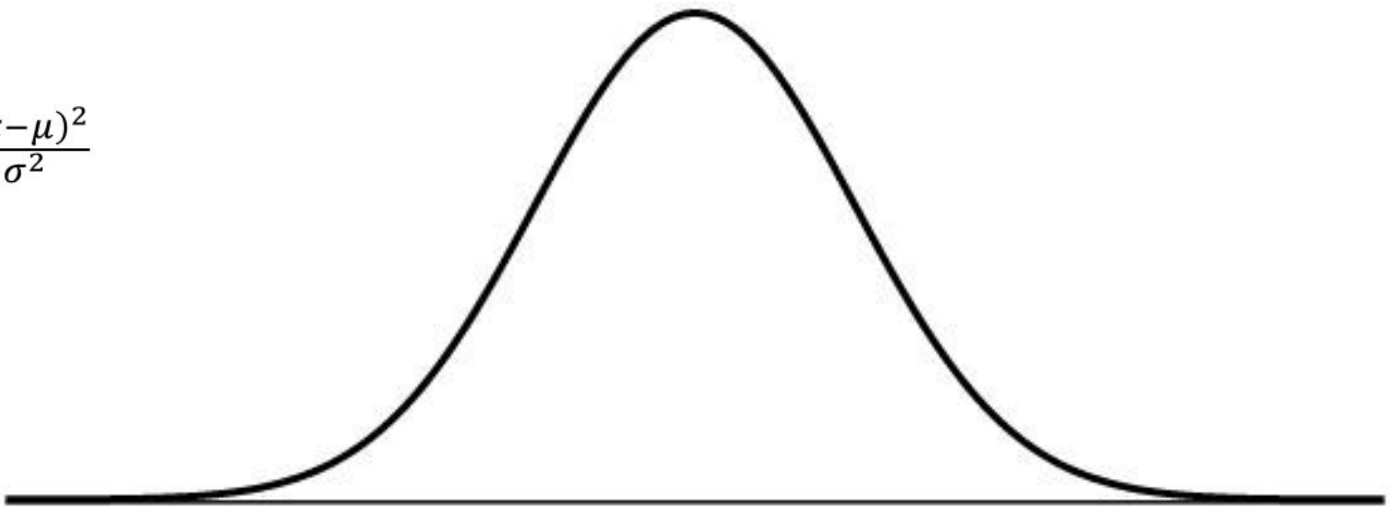
Exponential Distribution

$$f(x) = \begin{cases} \lambda e^{-\lambda * x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



Normal Distribution

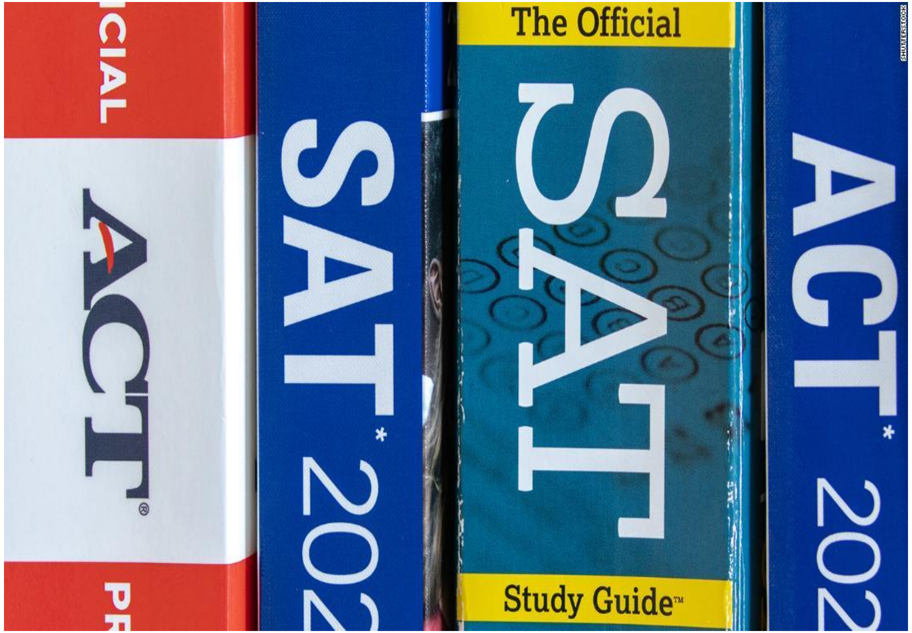
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

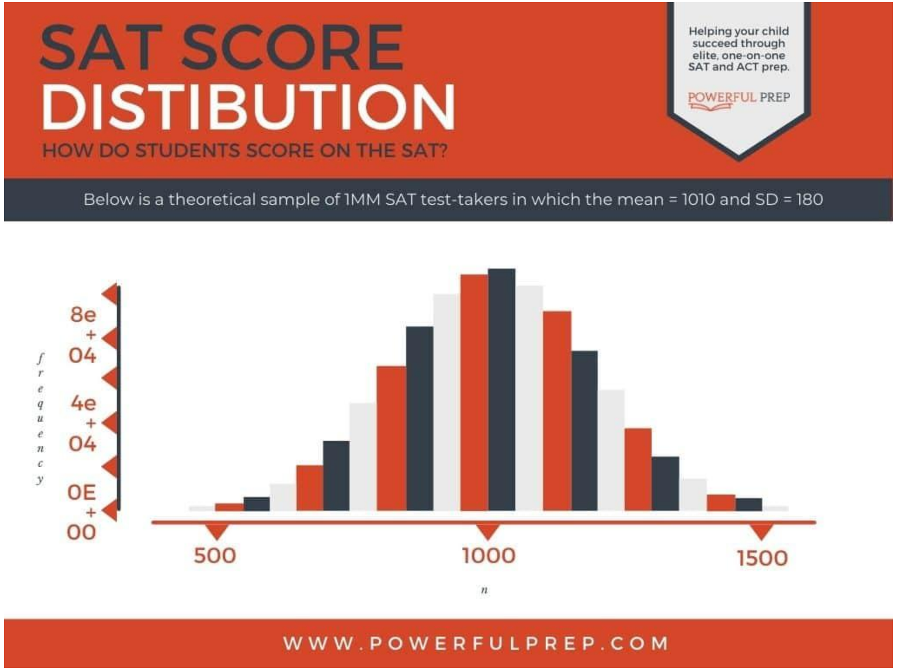


Normal Distribution

Normal Distribution

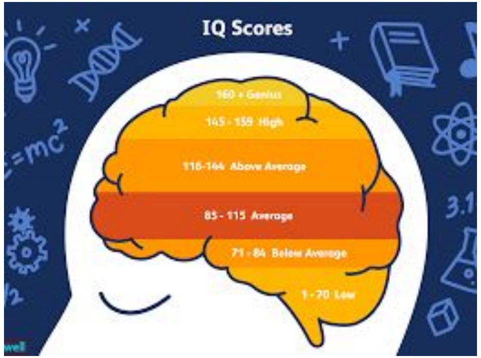
Normal Distribution





Normal Distribution

Normal Distribution



Interesting Case: binomial becomes normal

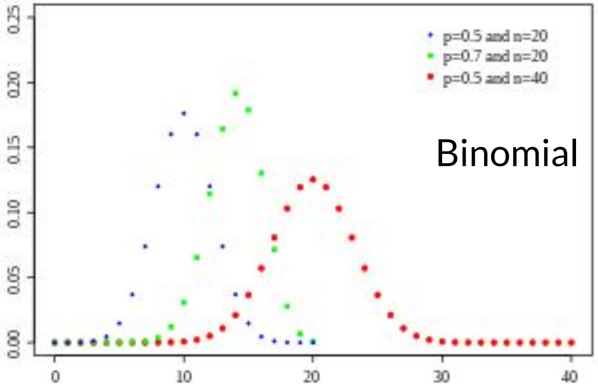
Binomial: $\binom{n}{k} p^k q^{n-k}$

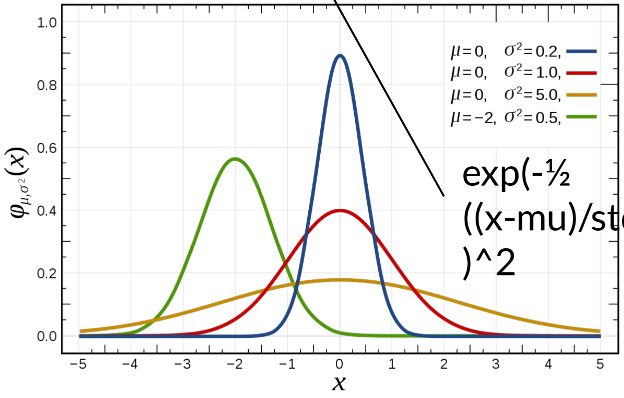
De Moivre-Laplace Theorem

Let X be a binomial random variable with parameters n and p. Then for any numbers a and b, a < b

$$\lim_{n\to\infty} P\left(a < \frac{X - np}{\sqrt{np(1-p)}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\, dt.$$

Note that np and $\sqrt{np(1-p)}$ appearing in this formula are, respectively, $E(X)$ and σ_X.





exp(-½ ((x-mu)/stdev)^2

Motivation for generalization

Let $X_1,\ X_2,\ X_3,\ \ldots$ be a sequence of independent Bernoulli random variables each with parameter p. Then $\forall i,\ E(X_i) = p$, $\mathrm{Var}(X_i) = p(1-p)$, and

$$\lim_{n\to\infty} P\Big(\frac{X_1 + X_2 + \cdots + X_n - np}{\sqrt{np(1-p)}} \le x\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy.$$

Chebyshev, 1887

Markov,

Lyapunov, 1901

The Central Limit Theorem

$$\lim_{n\to\infty} P(Z_n \le x) = \lim_{n\to\infty} P\Big(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le x\Big)$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy.$$

Remark 11.2 Note that Z_n is simply $X_1 + X_2 + \cdots + X_n$ standardized. ♦

Remark 11.3 In Theorem 11.12, let $\bar{X}$ be the mean of the random variables X_1, X_2, . . ., X_n. The central limit theorem is equivalent to

$$\lim_{n\to\infty} P\Big(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \le x\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy.$$

It is this version of the central limit theorem that has enormous applications in statistics.

♦