* Random Variable 随机变量

X 大写字母，是函数 (function)

可随意取值

$$X = \begin{cases} \boxed{21}, & \text{rain tomorrow} \\ \boxed{100}, & \text{No Rain} \end{cases}$$

不是变量，更像是函数

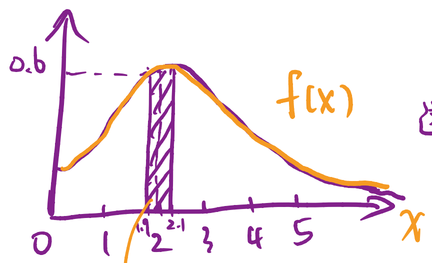
maps us from the world of a random process to a number

Can Take on either number, so Random

- **Discrete** random variables
  离散，有限个结果
- **Continuous** random variables
  连续，无限个结果

X = # facing up on a fair dice

Y = exact amount of rain tomorrow



$P(Y=2) = 0.5$?

这个2没有办法非常精确，直线量不了面积

$\therefore P(|Y-2|<0.1) = P(1.9<Y<2.1)$ = 阴影面积

可以问范围内的P，但无法得知一个点上的精确 number

面积 = 概率密度函数两点间的定积分

definite integral of this probability density function

$\int_{1.9}^{2.1} f(x)dx$

$\int_0^\infty f(x)dx = 1$ 同样适用于离散概率分布

permutation & combination
排列 組合

bionomial distribution

X = # of successes with probability p after n trials

$E(X) = n \cdot p$ choose k of n

期望

$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ The probability of get k chances

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \underbrace{0 \binom{n}{0} p^0 (1-p)^{n-0}}_{0} + \boxed{1 \binom{n}{1} p^1 (1-p)^{n-1} + \cdots + n \binom{n}{n} p^n (1-p)^{n-n}}$$

$$= \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$k!$ = k factorial

$$= \sum_{k=1}^{n} k \cdot \frac{n!}{\underbrace{\boxed{k!}}_{k(k-1)!}(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{\overbrace{\boxed{n!}}^{n(n-1)!}}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}$$

Make $a = k-1$, $b = n-1$ $\Rightarrow$ $a+1 = k$, $b+1 = n$ $\Rightarrow$ $n-k = b-a$

$$= np \boxed{\sum_{a=0}^{b} \frac{b!}{a!(b-a)!} p^a (1-p)^{b-a}} = \sum_{a=0}^{b} \binom{b}{a} p^a (1-p)^{b-a} = 1$$

*Poisson Distribution

泊松分布

From binomial distribution

$E(x) = \lambda = np$

① $\lim\limits_{x \to \infty} \left(1 + \frac{a}{x}\right)^x = e^a$

令 $\frac{1}{n} = \frac{a}{x}$ $\quad x = na$

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^{na} = \lim_{n\to\infty}\left(\left(1+\frac{1}{n}\right)^n\right)^a = \Big(\underbrace{\boxed{\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n}}_{e}\Big)^a = e^a$$

② $$\frac{x!}{(x-k)!} = \frac{x}{1}\frac{(x-1)}{2}\frac{(x-2)}{3}\cdots\frac{(x-k+1)}{k+2}$$

$E(x) = \lambda = np \qquad p = \frac{\lambda}{n}$

$$p(X=k) = \lim_{n \to \infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$\lim_{n\to\infty} \frac{n!}{(n-k)!\,k!} \cdot \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$\lim_{n\to\infty} \frac{n\cdot(n-1)(n-2)\cdots(n-k+1)}{n^k} \cdot \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$\because \lim\limits_{x\to a} f(x)g(x) = \lim\limits_{x\to a} f(x) \lim\limits_{x\to a} g(x)$

$$\therefore \lim_{n\to\infty} \frac{n^k + \cdots}{n^k} \cdot \frac{\lambda^k}{k!} \cdot \boxed{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n} \left(1-\frac{\lambda}{n}\right)^{-k} \qquad \left(1 - \frac{\lambda}{n}\right)^{-k}, \ \frac{\lambda}{n} \to 0$$

最高次项

$$= 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1$$

$$\therefore P(X=k) = \lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

*Law of Large Numbers 大数定律

$X,\ E(X)$

$$\widetilde{X_n} = \frac{x_1 + x_2 + \cdots + x_n}{n} \to E(x) \qquad \bar{X} \to \mu \quad \text{for } n \to \infty$$

样本越大.

试验做的次数越多，样本均值越接近期望值

*Normal distribution, the Gaussian distribution, the bell curve

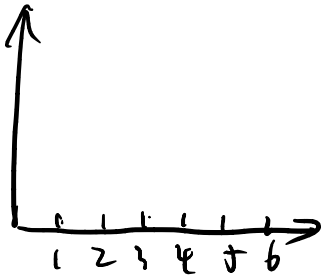
正态分布　　高斯分布　　钟形曲线

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

standard

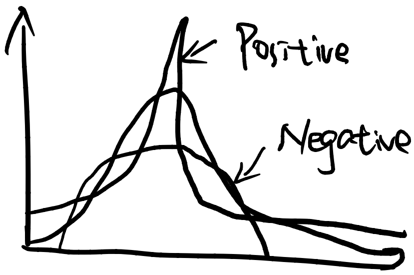
Central Limit Theorem
中心极限定理

正态分布



Sampling Distribution of the sample mean
样本均值的抽样分布

Kurtosis 峰度



sample mean
$n \to \infty$ → normal distribution
sample size

$$\sigma_{\bar{x}}^{2} = \frac{\sigma^{2}}{n}$$

Standard Error of the mean

$$\sigma_{\bar{x}}^{2} = \frac{\sigma}{\sqrt{n}}$$

Bernoulli Distribution

伯努力分布

success $= p$ failure $= 1-p$

$\mu = (1-p)\cdot 0 + p\cdot 1 = p$

$\sigma^2 = (1-p)(0-p)^2 + p(1-p)^2 = p(1-p)$

$\sigma = \sqrt{p(1-p)}$

置信区间

Find an interval such that "reasonably confident" that there is 95% chance that the true $\mu - p = \mu_{\bar{x}}$ is in that interval

$P(\bar{x}$ is within $2\sigma_{\bar{x}}$ of $\mu_{\bar{x}}) = 95.4\%$

$P(\mu_{\bar{x}}$ is within $2\sigma_{\bar{x}}$ of $\bar{x}) = 95.4\%$

$P(p$ is within $2(0.05)$ of $\bar{x}) = 95.4\%$

$P(p$ is within 0.10 of $\bar{x}) \approx 95\%$

$P(p$ is within $0.43 \pm 0.1) \approx 95\%$

$\therefore$ 95% confidence interval of 33% to 53%