数字媒体技术基础

Meng Yang

www. smartllv. com

SUN YAT-SEN University

机器智能与先进计算教育部重点实验室

智能视觉语言学习研究组

4 媒体信息表示基础

Course Outline

4 媒体信息表示基础 8

4.1 数字化过程8

4.1.1 采样的奈圭斯特理论 8

4.1.2 量化的信噪比理论 8

4.2 信息论基础8

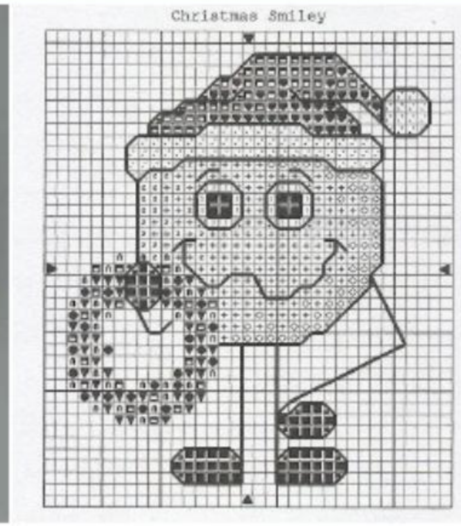
4.2.1 概率基础8

4.2.2 信息熵理论 8

4. 1数字化过程

数字化过程

- 图像数字化过程



由许多点组成的点阵图，我们称为位图（bitmap），构成位图的点称为像素（Pixel）。位图与我们生活中的手工“十字绣”很相似。

数字化过程

- 声音数字化

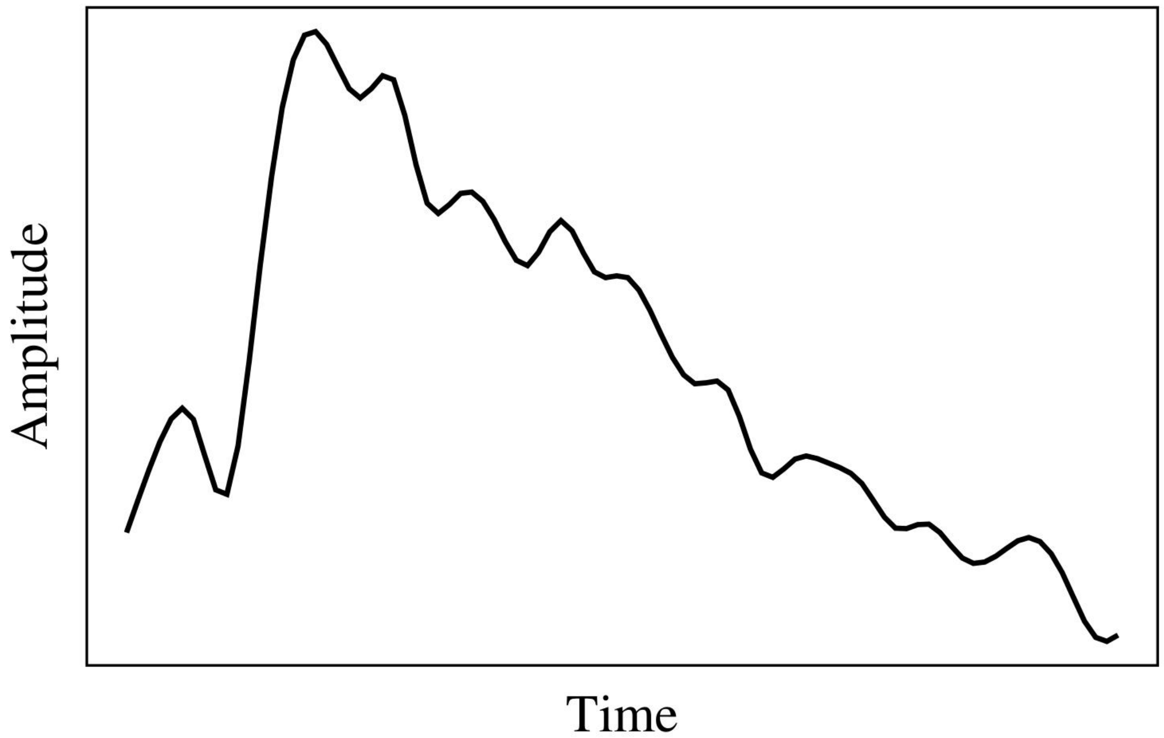


Fig. 6.1: An analog signal: continuous measurement of pressure wave.

数字化过程

- 声音数字化

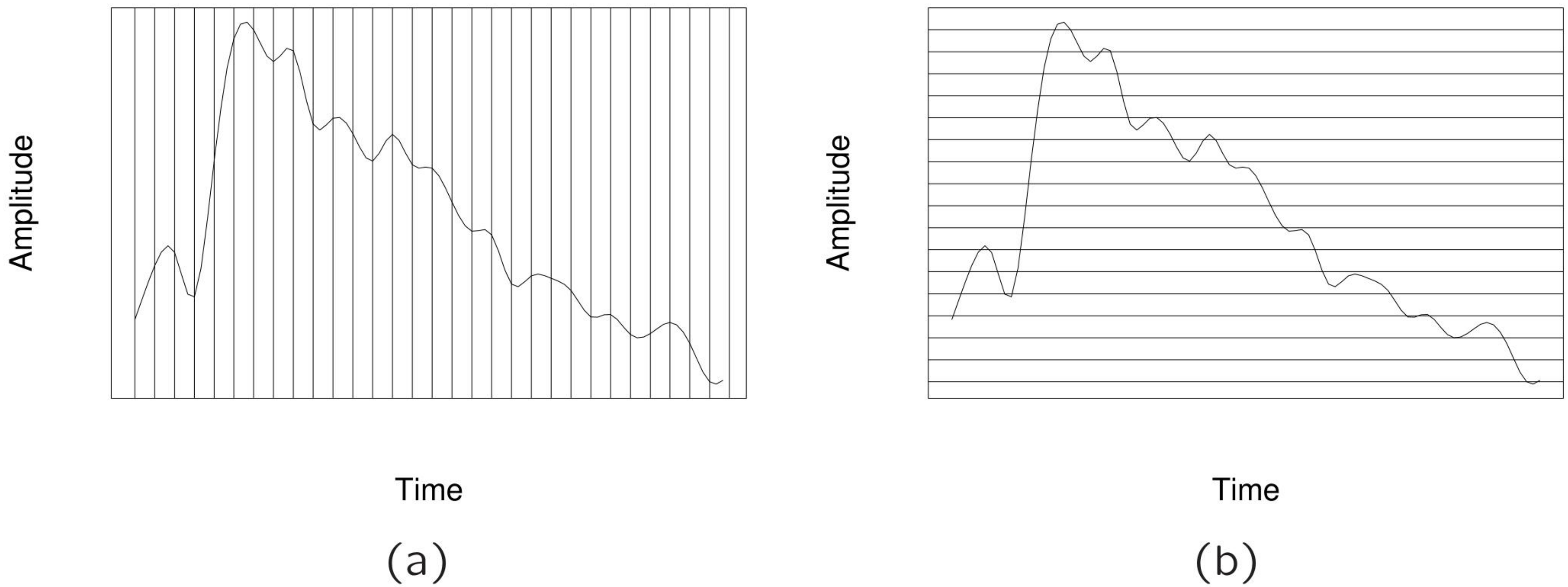


Fig. 6.2: Sampling and Quantization. (a): Sampling the analog signal in the time dimension. (b): Quantization is sampling the analog signal in the amplitude dimension.

数字化过程

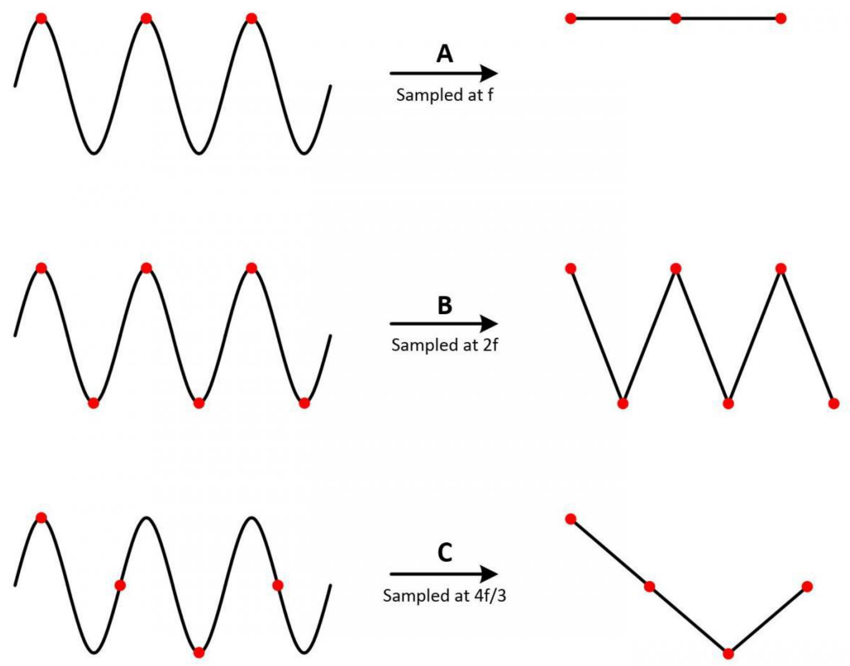
❑ 声音数字化

- Thus to decide how to digitize audio data we need to answer the following questions:

1. What is the sampling rate?
2. How finely is the data to be quantized, and is quantization uniform?
3. How is audio data formatted? (file format)

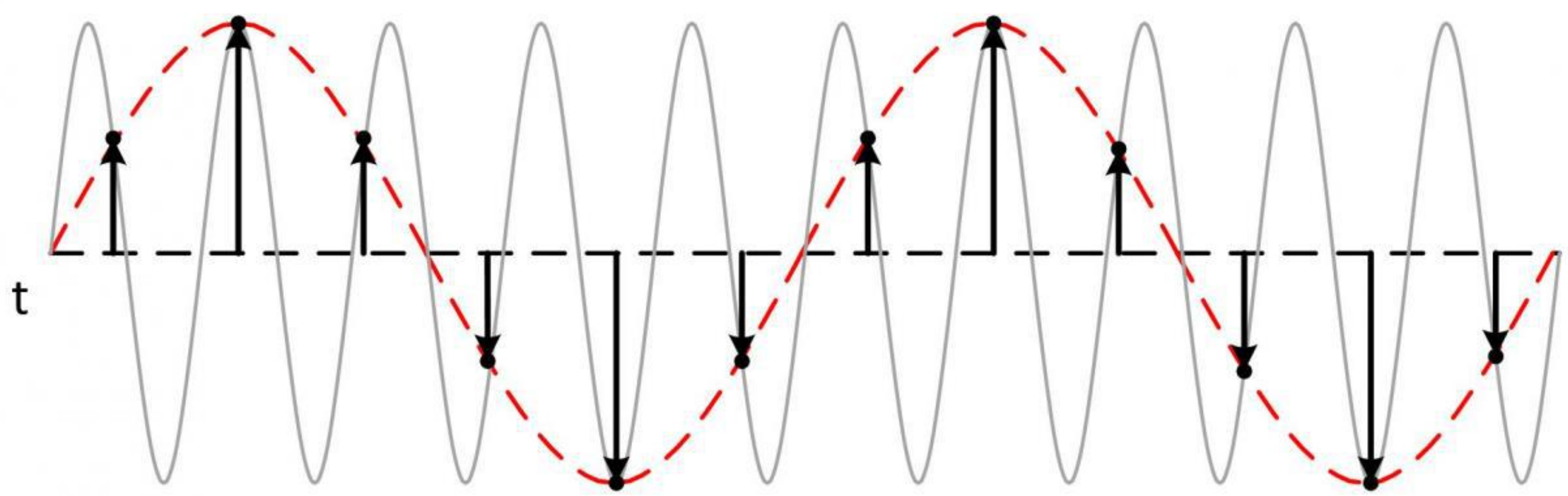
数字化过程

- 奈奎斯特采样定理



采样的奈圭斯特理论

- 混叠



采样的奈圭斯特理论

- 混叠

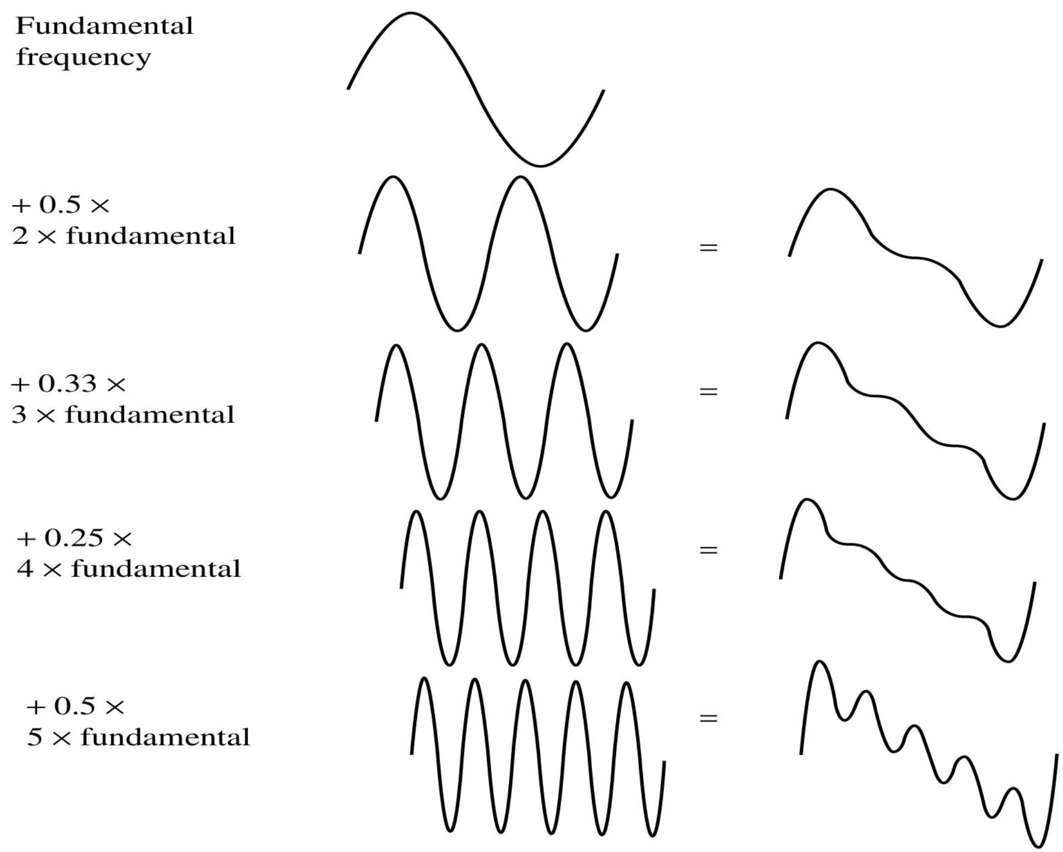


Fig. 6.3: Building up a complex signal by superposing sinusoids

采样的奈圭斯特理论

❑ 奈奎斯特采样定理

- The Nyquist theorem states how frequently we must sample in time to be able to recover the original sound.

  (a) Fig. 6.4(a) shows a single sinusoid: it is a single, pure, frequency (only electronic instruments can create such sounds).

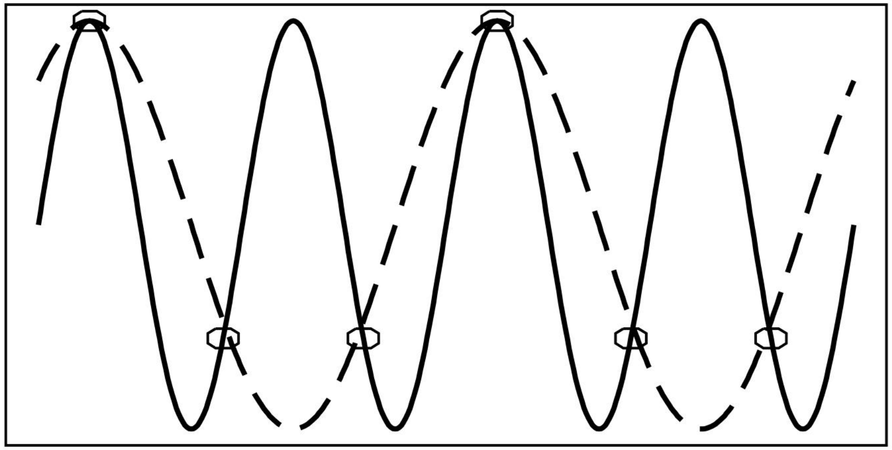
  (b) If sampling rate just equals the actual frequency, Fig. 6.4(b) shows that a false signal is detected: it is simply a constant, with zero frequency.

  (c) Now if sample at 1.5 times the actual frequency, Fig. 6.4(c) shows that we obtain an incorrect (**alias**) frequency that is lower than the correct one — it is half the correct one (the wavelength, from peak to peak, is double that of the actual signal).

  (d) Thus for correct sampling we must use a sampling rate equal to at least *twice the maximum frequency* content in the signal. This rate is called the **Nyquist rate**.

采样的奈圭斯特理论

❑ 奈奎斯特采样定理



(c)

Fig. 6.4: Aliasing. (a): A single frequency. (b): Sampling at exactly the frequency produces a constant. (c): Sampling at 1.5 times per cycle produces an *alias* perceived frequency.

采样的奈圭斯特理论

奈奎斯特采样定理

- **Nyquist Theorem**: If a signal is **band-limited**, i.e., there is a lower limit f_1 and an upper limit f_2 of frequency components in the signal, then the sampling rate should be at least $2(f_2 - f_1)$.

- **Nyquist frequency**: half of the Nyquist rate.
  - Since it would be impossible to recover frequencies higher than Nyquist frequency in any event, most systems have an **antialiasing filter** that restricts the frequency content in the input to the sampler to a range at or below Nyquist frequency. 反混淆

- The relationship among the Sampling Frequency, True Frequency, and the Alias Frequency is as follows:

$$f_{alias} = f_{sampling} - f_{true}, \quad \text{for} \quad f_{true} < f_{sampling} < 2 \times f_{true} \qquad (6.1)$$

量化的信噪比理论

量化信噪比理论

Signal to Noise Ratio (SNR)

- The ratio of the power of the correct signal and the noise is called the *signal to noise ratio* (**SNR**) — a measure of the quality of the signal.
- The SNR is usually measured in decibels (**dB**), where 1 dB is a tenth of a **bel**. The SNR value, in units of dB, is defined in terms of base-10 logarithms of squared voltages, as follows:

$$SNR = 10 \log_{10} \frac{V_{signal}^2}{V_{noise}^2} = 20 \log_{10} \frac{V_{signal}}{V_{noise}} \qquad (6.2)$$

量化的信噪比理论

❑ 量化信噪比理论

a) The power in a signal is proportional to the square of the voltage. For example, if the signal voltage V_{signal} is 10 times the noise, then the SNR is $20 * \log_{10}(10) = 20\text{dB}$.

b) In terms of power, if the power from ten violins is ten times that from one violin playing, then the ratio of power is 10dB, or 1B.

c) *To know:* Power — 10; Signal Voltage — 20.

量化的信噪比理论

量化信噪比理论

- The usual levels of sound we hear around us are described in terms of decibels, as a ratio to the quietest sound we are capable of hearing. Table 6.1 shows approximate levels for these sounds.

Table 6.1: Magnitude levels of common sounds, in decibels

Threshold of hearing	0
Rustle of leaves	10
Very quiet room	20
Average room	40
Conversation	60
Busy street	70
Loud radio	80
Train through station	90
Riveter	100
Threshold of discomfort	120
Threshold of pain	140
Damage to ear drum	160

量化的信噪比理论

量化信噪比理论

Signal to Quantization Noise Ratio (SQNR)

- Aside from any noise that may have been present in the original analog signal, there is also an additional error that results from quantization.

(a) If voltages are actually in 0 to 1 but we have only 8 bits in which to store values, then effectively we force all continuous values of voltage into only 256 different values.

(b) This introduces a roundoff error. It is not really "noise". Nevertheless it is called **quantization noise** (or quantization error).

量化的信噪比理论

量化信噪比理论

- The quality of the quantization is characterized by the Signal to Quantization Noise Ratio (**SQNR**).

  (a) **Quantization noise**: the difference between the actual value of the analog signal, for the particular sampling time, and the nearest quantization interval value.

  (b) At most, this error can be as much as half of the interval.

量化的信噪比理论

(c) For a quantization accuracy of N bits per sample, the SQNR can be simply expressed:

$$SQNR = 20 \log_{10} \frac{V_{signal}}{V_{quan\_noise}} = 20 \log_{10} \frac{2^{N-1}}{\frac{1}{2}}$$

$$= 20 \times N \times \log 2 = 6.02\, N (\mathrm{dB}) \qquad (6.3)$$

- Notes:

(a) We map the maximum signal to $2^{N-1} - 1$ ($\simeq 2^{N-1}$) and the most negative signal to -2^{N-1} .

(b) Eq. (6.3) is the *Peak* signal-to-noise ratio, PSQNR: peak signal and peak noise.

量化的信噪比理论

(c) The *dynamic range* is the ratio of maximum to minimum absolute values of the signal: V_{max}/V_{min}. The max abs. value V_{max} gets mapped to $2^{N-1}-1$; the min abs. value V_{min} gets mapped to 1. V_{min} is the smallest positive voltage that is not masked by noise. The most negative signal, $-V_{max}$, is mapped to -2^{N-1}.

(d) The quantization interval is $\Delta V=(2V_{max})/2^N$, since there are 2^N intervals. The whole range V_{max} down to $(V_{max}-\Delta V/2)$ is mapped to $2^{N-1}-1$.

(e) The maximum noise, in terms of actual voltages, is half the quantization interval: $\Delta V/2=V_{max}/2^N$.

量化的信噪比理论

- $6.02N$ **is the worst case**. If the input signal is sinusoidal, the quantization error is statistically independent, and its magnitude is uniformly distributed between 0 and half of the interval, then it can be shown that the expression for the SQNR becomes:

$$SQNR = 6.02N + 1.76(dB) \tag{6.4}$$

目 录

4.2信息论基础

信息论基础

- 概率基础——概率公理与随机变量
  - 概述
    - **随机试验(Random Experiment):**对随机现象做出的观察与科学实验。
    - **样本空间(Samplespace):**随机试验所有的基本可能结果构成的集合称Q。Q的元素为样本点(Sample point)。
    - **事件(Event)**是试验中“人们感兴趣的结果”构成的集合,是Q的子集。各种不同的事件的总体构成——个 事件集合，称为事件域F

概率公理与随机变量

概述

- 事件是随机的。赋予事件一个出现可能性的度量值，称为概率(Probability)。

- “可能性的度量值”是“宏观”意义下(即大数量的情形下)的比例值，由相对频率(Relative frequency)来计算,

$$P(A) \approx \frac{\text{试验中}A\text{出现的次数}}{\text{总试验次数}} = \frac{n_A}{n} \qquad (n\text{很大})$$

概率基础

- 概率公理与随机变量
  - 概率公理:任何事件A的概率满足:
    - (1)非负性:任取事件A，$P(A)\geq 0$
    - (2)归一性:$P(\Omega)=1$
    - (3)可加性:若事件A,B互斥，即，$A\cap B=\varnothing$,则，$P(A\cup B)=P(A)+P(B)$

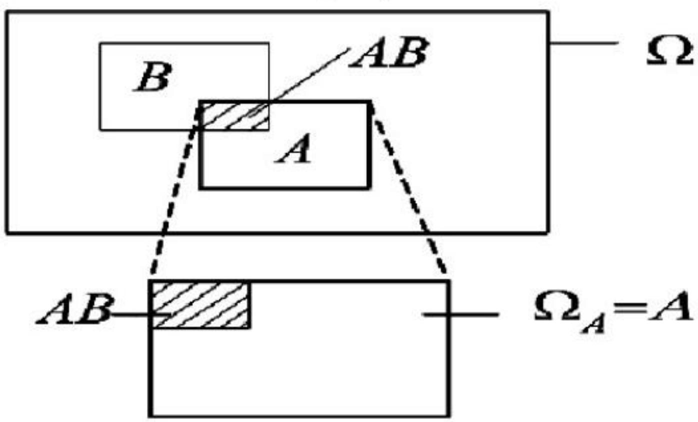
概率基础

- 概率公理与随机变量
  - 事件概率的基本性质:
    - (1) $P(\varnothing)=0$
    - (2) $0\leqslant P(A)\leqslant 1$
    - (3) $P(A)\leqslant P(B)$，如果$A\subseteq B$
    - (4) $P(AB)\leqslant P(A)\leqslant P(A\cup B)$

概率基础

- 概率公理与随机变量
  - **条件事件：** B|A = 事件A发生条件下的事件B
  - **条件概率** （Conditional probability）

$$P(B|A)=\frac{P(AB)}{P(A)}, \quad P(A)>0$$



概率基础

概率公理与随机变量

- 事件A与B独立（Independent）等价地定义为

$$P(AB)=P(A)P(B)$$

- 多个事件$A_1, A_2, \cdots, A_n$彼此独立，

$$P\left(A_{k_1}A_{k_2}\cdots A_{k_m}\right)=P\left(A_{k_1}\right)P\left(A_{k_2}\right)\cdots P\left(A_{k_m}\right)$$

概率公理与随机变量

例1.1 分析掷均匀硬币问题。

解： H – 正面，T – 反面。因此，

（1） 样本空间： $\Omega=\{H,T\}$

（2） 事件域： $F=\{\{H\},\{T\},\varnothing,\Omega\}$

（3） 由硬币的均匀特性可得，

$P\{H\}=P\{T\}=0.5$；而且 $P\{\varnothing\}=0$，$P\{\Omega\}=1$

概率公理与随机变量

例1.2 一列N个格子，将一只小球随机放入其中任一格子。求：（1）小球放入第k号格子的概率？（2）前k个格子中有小球的概率？

解： 因为是等概的，显然，

$$P(\text{小球放入任一格子})=\frac{1}{N}$$

又各个格子是互斥的，于是，

$$P(\text{小球放入任意}k\text{个格子})=\frac{k}{N}$$

概率基础

- 概率公理与随机变量
  - 随机变量

在样本空间 Ω 上定义一个单值实函数 $X(\xi)$，称为随机变量（Random variable，常缩写为r.v.）。并规定：用 $\{X(\xi)\le x\}$ 的概率来描述 $X(\xi)$ 的概率特性，记为

$$F_X(x)=P\{X(\xi)\le x\}$$

称它为 X 的分布函数（Distribution function），或称为累积分布函数（Cumulative distribution function）。

概率公理与随机变量

随机变量的类型：

1. **连续型**：$F(x)$ 是连续取值的。易见，

$$P(X = x) = 0$$

2. **离散型**：$F(x)$ 仅含有跳跃型间断点：$\{x_i\}$；仅在这些点上有非零的概率：$\{p_i\}$，

$$P(X = x_i) = p_i \quad (i\text{为整数})$$

称为 X 的分布律（或分布列）（Distribution law）。

3. **混合型**：上面两种形式的组合。

❑ 概率公理与随机变量

X 概率密度函数 (Probability density function)

$$f(x)=\frac{d}{dx}F(x)$$

基本性质为:

1. $f(x)\geq 0, \qquad \int_{-\infty}^{+\infty} f(x)\,dx=1$

2. $P\{X\in A\}=\int_A f(x)\,dx$

概率公理与随机变量

对于分布律为$P(X=x_i)=p_i$的离散型随机变量，其分布函数形如：

$$F(x)=\sum_i p_i u(x-x_i),\quad (i\text{为整数})$$

密度函数为

$$f(x)=\sum_i p_i \delta(x-x_i),\quad (i\text{为整数})$$

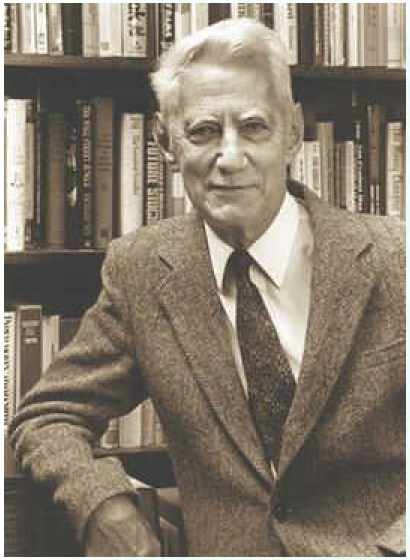
式中，取值位置对应$u(\ )$与$\delta(\ )$自变量的偏移量，取值概率对应前面的幅值。

概率公理与随机变量

- 随机变量不同于普通变量表现在两点上：
  - (1) 变量可以有多个取值，并且永远不能预知它到底会取哪个值；
  - (2) 变量取值是有规律的，这种规律用概率特性来明确表述；

信息熵理论

- 信息论之父
  - 克劳德·艾尔伍德·**香农**



说个香农的，他有许多爱好，特别令人难以置信的是香农可以熟练地玩一套杂技。不是在舞台上，而是在日常生活中，例如在贝尔实验室的走廊里。

从MIT到香农宽敞的住宅只有几英里。他的住宅里放满了各种乐器，诸如有5台钢琴、30多种其他乐器，从短笛到各种铜管乐器应有尽有。童年时代，他热衷于装无线电收音机、练莫尔斯电报码、搞密码学等。在Gaylord 上中学时他还当过Western Union 的信使。

信息熵理论

信息量

在数学上，设事件A的概率是P(A)，则称 $H(A) = -\log_2 P(A)$ 为A带来的信息量。

例如，抛一次硬币，它带来的信息量是 $-\log_2\left(\frac{1}{2}\right)$ =1，抛两枚硬币，它带来的信息量是 $-\log_2\left(\frac{1}{4}\right)$ =2，抛两枚硬币一共有4种可能的结果，每种结果的概率是1/4。

信息熵

- The *entropy* η of an information *source* with alphabet $S = \{s_1, s_2, \ldots, s_n\}$ is:

$$\eta = H(S) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} \tag{7.2}$$

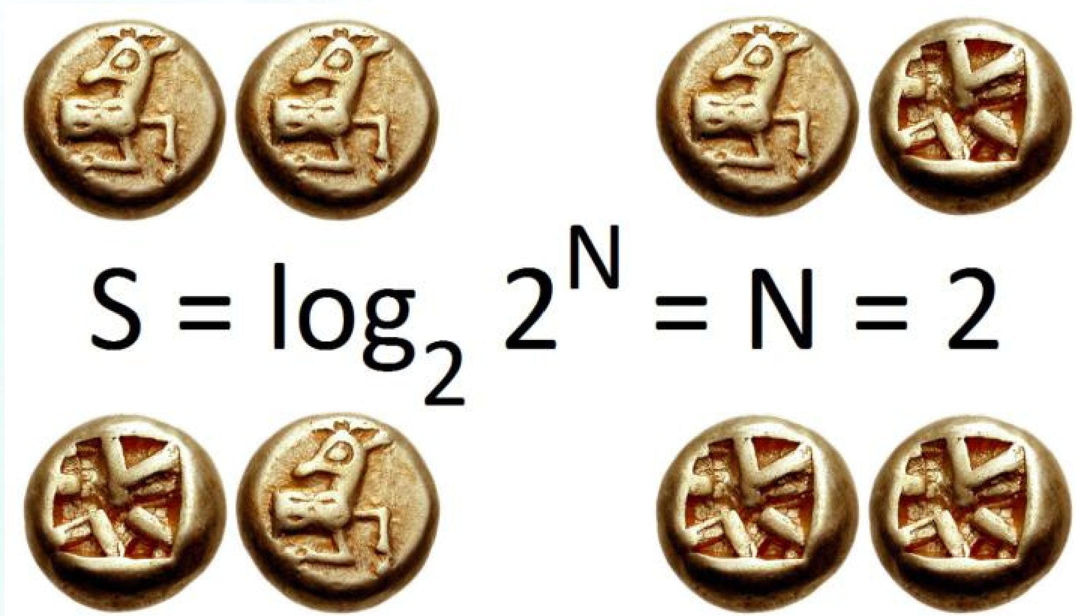
$$= -\sum_{i=1}^{n} p_i \log_2 p_i \tag{7.3}$$

p_i – probability that symbol s_i will occur in S.

$\log_2 \frac{1}{p_i}$ – indicates the amount of information (*self-information* as defined by Shannon) contained in s_i, which corresponds to the number of bits needed to encode s_i.

信息熵理论

- 信息熵



信息熵理论

❑ 世界杯信息量

举个吴军在《数学之美》中一样的例子，假设世界杯决赛圈32强已经产生，那么随机变量“2018年俄罗斯世界杯足球赛32强中，谁是世界杯冠军？”的信息量是多少呢？