Probability Theory

- ❑Probabilistic Concepts
- ❑Random Variables
- ❑Statistical Averages

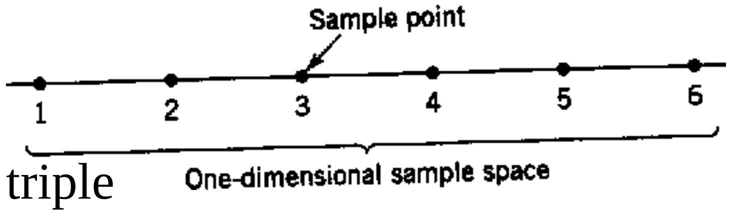
Probabilistic Concepts (1/4)

- ❑前言
  - ❁機率理論根植於實驗結果受制於機會的現象
    - ☞因潛在隨機現象或機會機制，實驗結果可能不同。此即隨機實驗，如丟銅板
    - ☞隨機實驗的三個特質
      - ✐在相同條件下可以重複
      - ✐在每次實驗的試驗中，其結果是不可預知的
      - ✐大量的實驗試驗，其結果具統計規律性，即實驗重複夠多次則可觀察到結果有一明確平均型態
- ❑Relative-frequency approach
  - ❁事件A的相對頻率：$N_n(A)/n$
    - ☞事件A：隨機實驗的一種可能結果
    - ☞$N_n(A)$：在n試驗中，事件A出現的次數
    - ☞相對頻率為小於等於1的非負實數

Probabilistic Concepts (2/4)

- 可定義事件A的機率為 $P(A)=\lim_{n\to\infty}\left(\frac{N_n(A)}{n}\right)$
  - 實驗若具統計規律性則當n夠大時，任何序列中A之相對頻率收斂至相同極限
  - 事件的機率為其在長試驗序列中所佔次數的比例
  - 代表實驗之嘗試中事件將出現的可能性
  - 某些應用不適合，如股市，因為難重複
  - 另一種適合的方式為將機率測量應有的特性視為原理，以相對頻率解釋來證明
- Axioms of probability
  - 術語
    - 取樣點：實驗的第k個(非第k次)結果，表為s_k(如H或T)
    - 取樣空間：實驗所有可能結果的集合體，表為S(如{H,T})
    - 事件：單一或一組取樣點。S為確定事件，空集合ϕ為空或不可能事件，為s_k基礎事件

Probabilistic Concepts (3/4)



- 組成機率系統的triple
  - 基礎事件(結果)的取樣空間S
  - 事件(即S子集合)之類別$\mathcal{E}$(可想成某種分類方式)
  - 一機率測量P(．)給在$\mathcal{E}$中的每個A，具機率原理
    - P(S)=1：確定事件之機率為1。
    - 0≤P(A)≤1：一事件之機率為非負實數，小於等於1
    - 相互排外事件聯集A+B，則P(A+B)=P(A)+P(B)
- 一些機率基本特性
  - $P(\bar{A})=1-P(A)$，由 $S=A+\bar{A}$ 證明
  - 若$A_1+A_2+\ldots+A_M=S$，則$P(A_1)+P(A_2)+\ldots+P(A_M)=1$，由原理一及三證明。若所有事件相等可能的，則$P(A_i)=1/M$

Probabilistic Concepts (4/4)

☞P(A+B)=P(A)+P(B)-P(AB)，P(AB)為聯合事件“A且B”的機率，稱聯合機率。原理三為其特例

❑Conditional Probability

- 給定A發生下B的機率P(B|A)=P(AB)/P(A)
  - ☞相對頻率 $P(B\mid A)=\lim_{n\to\infty}(\frac{N_n(AB)}{N_n(A)})=\lim_{n\to\infty}(\frac{N_n(AB)/n}{N_n(A)/n})=\frac{P(AB)}{P(A)}$
  - ☞P(AB)=P(B|A)P(A), P(AB)=P(A|B)P(B)
  - ☞Bayes' rule: P(B|A)=P(A|B)P(B)/P(A)
  - ☞統計上獨立的：給定A下B的機率同B的基礎機率
    - P(B|A)=P(B) → P(AB)=P(A)P(B) → P(A|B)=P(A)

Random Variables (1/2)

❑隨機變數：一個函數其定義域為取樣空間(如H或T)，值域為實數的某個集合(如+1或-1)

- s為實驗結果，則隨機變數表為X(s)或僅X
  - ☞離散隨機變數：如丟骰子
  - ☞連續隨機變數：雜訊電壓
- (累積)分佈函數cdf $F_X(x)=P(X\le x)$
  - ☞是x的函數，非X的函數，但相依於X的分配
  - ☞有特性：介於0與1之間；非遞減函數
- 機率密度函數$f_X(x)=dF_X(x)/dx$
  - ☞ $P(x_1<X\le x_2)=\int_{x_1}^{x_2} f_X(x)dx, F_X(x)=\int_{-\infty}^{x} f_X(\xi)d\xi, \int_{-\infty}^{\infty} f_X(x)dx=1$
  - ☞非負函數(可能大於1)且全面積為1

Random Variables (2/2)

- 聯合分佈函數與聯合機率密度函數
  - $F_{X,Y}(x,y)=P(X\le x, Y\le y)$ 為非遞減函數
  - $f_{X,Y}(x,y)=\dfrac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$ 為非負函數，且全體積為1
  - 邊際密度 $f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,\eta)d\eta$ (微分 $F_X(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X,Y}(\xi,\eta)d\xi d\eta$)
- 給定X=x下Y的條件機率 $f_Y(y\,|\,x)=f_{X,Y}(x,y)/f_X(x)$
  - 為y的函數，x則為任意固定值
  - 統計上獨立的：X的結果不影響Y的分佈
    - $f_Y(y\,|\,x)=f_Y(y),\ f_{X,Y}(x,y)=f_X(x)f_Y(y)$
    - 聯合機率密度函數=各別邊際密度的乘積則統計上獨立

Statistical Averages (1/3)

- 前言
  - 期望值或平均：$\mu_X=E[X]=\int_{-\infty}^{\infty} x f_X(x)dx$
    - E為統計期望運算子，μ_X定位出機率密度曲線下面積之重心
    - 使用近似和之極限來說明：取樣平均

$$E[X]=\lim_{\Delta\to 0}\sum_{k=-\infty}^{\infty}\int_{k\Delta}^{(k+1)\Delta} x_k f_X(x)dx=\lim_{\Delta\to 0}\sum_{k=-\infty}^{\infty} x_k P(x_k-\tfrac{\Delta}{2}<X\le x_k+\tfrac{\Delta}{2})\approx\sum_{k=-\infty}^{\infty} x_k(\tfrac{N_n(k)}{n})=\frac{1}{n}\sum_{k=-\infty}^{\infty} x_k N_n(k)$$

每個x_k加權其發生次數後全部相加再除以全部觀察次數

  - Y=g(X)的期望值 $\int_{-\infty}^{\infty} y f_Y(y)dy=E[Y]=E[g(X)]=\int_{-\infty}^{\infty} g(x)f_X(x)dx$

Statistical Averages (2/3)

❑Moments

- 第n個機率分佈的動量 $E[X^n]=\int_{-\infty}^{\infty}x^n f_X(x)dx$
  - n=1即平均，n=2為均方值
- 中心動量：X-μ_X的動量 $E[(X-\mu_X)^n]=\int_{-\infty}^{\infty}(x-\mu_X)^n f_X(x)dx$
  - n=1為0，n=2為變異數var[X]=σ_X^2，σ_X為標準差
- σ_X^2有測量隨機變數之隨機性的意義
  - 給定σ_X^2，基本上限制了$f_X(x)$對應μ_X的有效寬度
  - Chebyshev不等式$P(|X-\mu_X|\geq\varepsilon)\leq\frac{\sigma_X^2}{\varepsilon^2}$，可部分描述機率分佈
  - 使用E的線性特性得$\sigma_X^2=E[(X-\mu_x)^2]=E[X^2]-\mu_X^2$
    - μ_X=0則 $\sigma_X^2=E[X^2]$

Statistical Averages (3/3)

❑Characteristic function $\phi_X(v)=E[\exp(jvX)]$

- 與Fourier transform相似，但差個負號
  - v對x與ω對t扮演相同角色，所以 $f_X(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_X(v)\exp(-jvx)dv$

❑Joint moments $E[X^iY^k]$: i=k=1為correlation

- Covariance cov[XY]=E[XY]- $\mu_X\mu_Y$
  - Correlation coefficient ρ=cov[XY]/ $\sigma_X\sigma_Y$
  - Uncorrelated iff cov[XY]=0
  - Orthogonal iff E[XY]=0
  - Given $\mu_X\mu_Y$=0, if E[XY]=0 then cov[XY]=0, 反之亦同(意義：說明不相關與正交之關係)
  - X, Y不相依，則不相關，反之不成立
    - Indep → $f_{X,Y}(x,y)=f_X(x)f_Y(y)$ → E[XY]=E[X]E[Y] → cov[XY]=0 → uncorrelated
    - uncorrelated → cov[XY]=0 → E[XY]=E[X]E[Y] → ?