概率论与数理统计

- HITwh 2191210307刘鑫龙
- 事件的运算及概率
  - 事件及运算
    - 德摩根律（长杠变短杠，开口换方向）
    - 加法公式
      - $P(A\cup B)=P(A)+P(B)-P(AB)$
        $P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$
    - 减法公式
    - 对立事件
      - $P(\bar{A})=1-P(A)$
    - 独立事件
  - 古典概型
    - 抽奖
  - 几何概型
    - 约会问题
- 全概率、贝叶斯公式
  - 条件概率
  - 乘法公式
  - 全概率公式
  - 贝叶斯公式
- 一维随机变量
  - 离散型随机变量
    - 分布律、分布函数
      - 分布函数的性质
        - $F(x)=P\{X\le x\}$
          ①$0\le F(x)\le 1$
          ②$F(-\infty)=0$， $F(+\infty)=1$
          ③$F(x)$ 单调不减函数
          ④ 右连续 $\lim_{x\to x_0^+}F(x)=F(x_0)$
    - 函数的分布
      - g=g(X)的分布律
  - 连续型随机变量
    - 概率密度、分布函数
      - 概率密度f(x)
        - 概率密度f(x)的性质
      - 分布函数
        - 分布函数F(x)的性质
    - 函数的分布
      - g=g(X)单调可导
        1. 求出 $g(X)$ 值域
        2. $x=h(y),x'=h'(y)$
        3. $f_Y(y)=f_X[h(y)]\cdot h'(y)$
      - g=g(X)非单调可导
        1. 求出 $g(X)$ 值域（根据定义域先写值域）
        2. $F_Y(y)=PY\le y=Pg(X)\le y=PX\le h(y)=F_X(h(y))$
        3. $F_Y'(y)=f_Y(y)$
- 五种重要分布
  - 离散型
    - 二项分布
    - 泊松分布
  - 连续型
    - 均匀分布
    - 指数分布
      - 指数函数的无记忆性
    - 正态分布
      - 标准正态分布
- 离散型 二维随机变量
  - 分布律
  - 边缘分布律
  - 独立性
  - 函数的分布
  - 条件分布
- 连续型二维随机变量
  - 分布函数
    - $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\,du\,dv$
  - 概率密度（联合概率密度）
    - $f(x,y)$
    - 性质
      - $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\,dx\,dy=1$
  - 边缘概率密度
    - 关于X
      - $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,dy$
    - 关于Y
      - $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\,dx$
  - 条件概率密度
    - 在Y=y的条件下，X的条件概率密度为
    - 在X=x的条件下，Y的条件概率密度为
  - 独立性
    - $F(x,y)=F_X(x)F_Y(y)\iff X$ 与 Y 相互独立
    - 要判断两个随机变量是否独立，只要验证X和Y的"联合分布"是否等于"边缘分布的乘积"就可以了。
- 连续型二维随机变量函数的分布
- 数学期望、方差、协方差
  - 一维随机变量期望与方差
    - 离散型
      - 数学期望
        - $EX=\sum_i x_ip_i$
      - 方差
    - 连续型
      - 数学期望
        - $EX=\int_{-\infty}^{+\infty}xp(x)\,dx$
      - 方差
        - $DX=EX^2-(EX)^2$
    - 常用分布的数学期望和方差
    - 期望E(X)
      - 性质
        - $EC=C$
          $E(CX)=CEX$
          $E(X_1+X_2+\cdots+X_n)=EX_1+EX_2+\cdots+EX_n$
          若 X_1， X_2， $\cdots$， X_n 相互独立，则 $E(X_1X_2\cdots X_n)=EX_1EX_2\cdots EX_n$
    - 方差D(X)
      - 性质
        - $DC=0$
          $D(CX)=C^2DX$
          若 X_1， X_2， $\cdots$， X_n 相互独立， $D(X_1+X_2+\cdots+X_n)=DX_1+DX_2+\cdots+DX_n$
          $DX=0$ 的充要条件是 X 取某一常数值 a 的概率为 1，即 $P\{X=a\}=1$（此处 a 就是 X 的数学期望）
  - 随机变量函数
    - 一维随机变量函数期望与方差
      - 离散型
        - $EY=Ef(x)=\sum_{i=1}^{\infty}f(x_i)p_i$
      - 连续型
        - $EY=Ef(X)=\int_{-\infty}^{+\infty}f(x)p_X(x)\,dx$
    - 二维随机变量期望与方差
      - 离散型
        - $EZ=Ef(X,Y)=\sum_{i=1}^{n}\sum_{j=1}^{n}f(x_i,y_j)p_{ij}$
      - 连续型
        - 求谁的数学期望就把谁放到被积函数里面去
        - $EZ=Ef(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)p(x,y)\,dx\,dy$
      - 由（X,Y）的概率密度p(x,y)求X,Y的数学期望公式
        - $EX=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\,dx\,dy$
        - $EY=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yp(x,y)\,dx\,dy$
  - 协方差
    - 基础知识
      - 方差
      - 协方差
      - 协方差性质
      - 相关系数
      - 随机变量X,Y不相关与下面的每一个结论等价
      - $\rho_{XY}=0$ 为 X 和 Y 不相关 (独立一定不相关，不相关不一定独立）
    - 矩，协方差矩阵
      - 数学期望是一阶原点矩
      - 方差是二阶中心矩
- 大数定理及中心极限定理
  - 切比雪夫不等式
    - 方差DX存在
      - $P\{|X-E(X)|\ge\varepsilon\}\le\frac{D(X)}{\varepsilon^2}$
      - $P\{|X-E(X)|<\varepsilon\}\ge 1-\frac{D(X)}{\varepsilon^2}$
  - 贝努利大数定理
    - n重贝努利试验，成功次数Y_n，每次实验成功概率p
  - 切比雪夫大数定理
    - 相互独立，DX_i ＜ C
  - 辛钦大数定理
    - 独立同分布，有限期望EX_i
  - 依概率收敛
  - 练习题做一下
- 抽样分布
  - 简单随机抽样
    - 独立同分布
  - 三种常见分布
    - χ^2 分布
    - t 分布
    - F 分布
  - 常用统计量及性质
    - 统计量不含任何未知参数
    - 常用统计量
      - 样本均值
        - $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$
      - 样本方差
        - $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n-1}(\sum_{i=1}^{n}X_i^2-n\bar{X}^2)$
      - 样本k阶原点矩
        - $a_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k,k=1,2,3\cdots$
      - 样本k阶中心矩
        - $A_k=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^k,k=1,2,\cdots$
    - 常考性质
      - $E(\bar{X})=\mu$
        - $D(\bar{X})=\frac{D(X)}{n}$
      - $E(S^2)=D(X)$
      - $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$
        - $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$
  - 抽样分布
    - 样本均值的分布
      - 定理6.4
        - $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$
      - 推论
        - $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$
    - 样本方差的分布
      - 定理6.5
        - $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$
    - 定理6.6
      - $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$
  - 三定理
    - 定理：设 X_1， X_2， ……，为总体 $N(\mu,\sigma^2)$ 的一个样本，则
      - $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$ 或 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$
      - $\bar{X}$ 与 S^2 相互独立，且 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$
      - $\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$
- 参数估计
  - 矩估计
    - $u_1=E(X)=f(\theta)$
    - $\theta=f^{-1}(u_1)$
    - $\hat{\theta}=f^{-1}(\bar{X})$
  - 最大似然估计
    - $L(\theta)=\prod_{i=1}^{n}f(x_i,\theta)$
    - 取对数 $lnL(\theta)=ln\prod_{i=1}^{n}f(x_i,\theta)$
    - 求导 $\frac{d}{d\theta}lnL(\theta)=0$
    - 解出 $\hat{\theta}=?$
  - 无偏估计
    - 若 $E(\triangle)=\square$，则 $\triangle$ 为 $\square$ 的无偏估计。
    - $E(\bar{X})=\mu$, 样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 为期望 μ 的无偏估计。
    - $E(S^2)=D(X)$，样本方差 S^2 为 $D(X)$ 的无偏估计。
    - $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$
- 区间估计
  - 置信度：1-α
  - 置信区间
    - 待估参数：μ
      - σ^2 已知
        - 选取随机变量 $u=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$
        - $P(-u_{\alpha/2}<u<u_{\alpha/2})=1-\alpha$
        - μ 置信区间为 $(\bar{x}-u_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{x}+u_{\alpha/2}\frac{\sigma}{\sqrt{n}})$
      - σ^2 未知
        - $t=\frac{\bar{x}-\mu}{s/\sqrt{n}}\sim t(n-1)$。
        - $P(-t_{\alpha/2}(n-1)<t<t_{\alpha/2}(n-1))=1-\alpha$
        - μ 置信区间为 $(\bar{x}-t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}},\bar{x}+t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}})$
    - 待估参数：σ^2
      - μ 未知
        - $\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$
        - $\chi^2_{1-\alpha/2}(n-1)<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha/2}(n-1)$
        - $(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)})$
  - 置信区间求解
    1. 定类型，摆公式
    2. 计算各分量
    3. 代入公式
    1. 选取随机变量
    2. 符合什么分布
    3. 写出置信区间
- 假设检验
  - 单个正态总体参数的显著性检验
    - 检验均值 μ
      - σ² 已知
        - 提出统计假设 $H_0:\mu=\mu_0$（显著性水平为 α）
        - 选取统计量 $u=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$，并根据样本值算出统计量 u 的值。
        - 查表求得满足 $P(-u_{\alpha/2}<u<u_{\alpha/2})=1-\alpha$ 的临界值 $u_{\alpha/2}$
        - 根据 $u=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ 构造小概率事件： $A=\{|u|\ge u_{\alpha/2}\}$
        - 下结论：如果 $|u|\ge u_{\alpha/2}$，则拒绝 H_0；否则，接受 H_0。
      - σ² 未知
        - 提出统计假设 $H_0:\mu=\mu_0$（显著性水平为 α）
        - 选取统计量 $t=\frac{\bar{x}-\mu}{s/\sqrt{n}}$，并根据样本值算出统计量 t 的值。
        - 查表求得满足 $P(-t_{\alpha/2}(n-1)<t<t_{\alpha/2}(n-1))=1-\alpha$ 的临界值 $t_{\alpha/2}$
        - 根据 $t=\frac{\bar{x}-\mu}{s/\sqrt{n}}\sim t(n-1)$ 构造小概率事件： $A=\{|t|\ge t_{\alpha/2}(n-1)\}$
        - 下结论：如果 $|t|\ge t_{\alpha/2}(n-1)$，则拒绝 H_0；否则，接受 H_0。
    - 检验方差 σ²
      - μ 未知
  - 两类错误
    - H_0 为真，否定了 H_0, 第一类错误：弃真，概率为 α
    - H_0 为假，接受了 H_0, 第二类错误：取伪，概率为 β
    - α 和 β 一个减少，另一个往往增大
- 近几年考点
  - 1
    - 分布函数的性质
    - 分布函数和概率密度的转换
    - 相关系数相关
    - 独立同分布相关
    - F分布
    - 方差公式
    - 置信区间，上侧分位数
    - 大题
      - 条件概率、边缘概率密度
      - 二维随机变量函数的分布函数与概率密度的求得
      - 矩估计量与最大似然估计量
  - 2
    - 各大分布的概率密度
    - 切比雪夫不等式
  - 3
    - 公式法通过边缘概率密度求边缘概率密度
  - 4
    - 分布函数的性质
  - 5
    - 均匀分布分布函数
    - t分布
    - Cov(X,Y)
    - 分布函数法
  - 6
    - 切比雪夫不等式
    - 置信区间
    - 左闭右开，右连续
    - 矩估计和极大似然估计
    - 子主题 5