

lec1

概念自查：

- statistics
- mean
 - conti
 - dis
- var
- sd
- iqr
- percentile
- rate
- z

lec2

统计相关

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- 条件概率
- 独立性如何表示
 - $P(A|B) = P(A)$
 - 反之
- 贝叶斯公式

随机变量

记rv为X

- cdf $F(x) = P(X \leq x)$ 有个等号
 - 判断是否为cdf的两个条件
 - ①值域0~1
 - ②单调非减
- 离散型rv
 - $\text{binom}(n, p)$
 - $E = np$
 - $\text{var} = np(1-p)$
 - multi $x \sim M(n, p_1 \dots p_k)$

$$p(x_1 \dots x_k) = \frac{1}{n!} \frac{1}{x_1! x_2! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}$$

- 泊松
 - 一段时间内
 - 参数: λ
- 连续型rv
 - normal
 - chi-square dist
 - n 个标准正态分布的平方的加和
 - t dist
 - 标准正态分布 $(X-\mu)/\sigma \sim N(0,1)$
 - t分布出现是因为不知道 σ (但知道 μ) , 用样本估计 σ^2
 - χ^2 分布
 - 星号部分 估计方差 σ^2

③ t_n $\frac{X-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \Rightarrow z\text{-score}$
 $\frac{\frac{\sigma}{\sqrt{n}}}{\sqrt{\frac{\chi_n^2}{n}}} \sim t_n$

④ $\Gamma(\alpha, \beta)$ $\frac{\chi_n^2}{n} \sim \Gamma\left(\frac{n}{2}, \frac{n}{2}\right)$
 $\frac{\frac{\chi_n^2}{n}}{\frac{\chi_m^2}{m}} \sim F_{n,m}$
 $\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \rightarrow \sigma^2$

CLT

构建渐进z score (? ? ? ?)

随机变量, 分布未知, 但 μ σ 已知

AIV

CLT \rightarrow Z-score

① $x_1 \dots x_n \sim f(\mu, \sigma^2)$

$$\frac{1}{n} E\left(\sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n E(x_i) = \mu$$

$$\text{Var}\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(x_i) = \frac{\sigma^2}{n}$$

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

estimation

在样本中可以知道的是 $\bar{x} \rightarrow E$; sample var 为了成为无偏估计, 分母是 $n-1$, 分子是 \bar{x} 不是 μ sample var $\rightarrow \sigma^2$

方法

- 点估计
- 区间估计

hypothesis testing

- H_0
- H_a
- 检验统计量
- 拒绝域
- 判断方式
 - 拒绝域
 - pvalue