概率论与数理统计

1 Chapter 5 大数定律与中心极限定理

切比雪夫不等式:

设随机变量X的数学期望和方差存在，分别记为μ,σ^2，则对任意的$\varepsilon > 0$, 有

$$P\{|X-\mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

辛钦大数定律:

设$\{X_i, i \geq 1\}$为独立同分布的随机变量序列，且数学期望存在，记为μ，则对任意的$\varepsilon > 0$，有

$$\lim_{n \to +\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| \geq \varepsilon\} = 0$$

辛钦大数定律推论:

设$\{X_i, i >= 1\}$为独立同分布的随机变量序列，若$h(x)$为一连续函数，且$E(h(X_i)) < +\infty$，则对任意的$\varepsilon > 0$,有

$$\lim_{n \to +\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} h(X_i) - a| \geq \varepsilon\} = 0$$

其中$a = E(h(X_i))$

2 Chapter 6 三大分布、四大定理

2.1 χ^2分布，t分布，F分布

2.1.1 （一）χ^2分布：

设$X_1, X_2, \ldots, X_n$为独立同分布的随机变量，且都服从标准正态分布$N(0,1)$
记：

$$Y = X_1^2 + X_2^2 + \ldots + X_n^2$$

则称Y服从自由度为n的χ^2分布，记为$Y \sim \chi^2(n)$

χ^2分布的性质

1. $\chi^2(2) = E(\frac{1}{2})$
2. 可加性：设$Y_1 \sim \chi^2(m), Y_2 \sim \chi^2(n)$，若两者相互独立，则$Y_1 + Y_2 \sim \chi^2(m+n)$
3. $E(\chi^2(m)) = m$,$Var(\chi^2(m)) = 2m$

2.1.2 （二）t分布：

设$X \sim N(0,1), Y \sim \chi^2(m)$，且$X, Y$相互独立
记：

$$t = \frac{X}{\sqrt{Y/m}}$$

则称t服从自由度为n的t分布，记为$t \sim t(n)$

t分布的性质

1. $f_t(x) = f_t(-x)$

2. $\lim\limits_{n \to +\infty} f_t(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

2.1.3 （三）F分布：

设$U \sim \chi^2(n_1), V \sim \chi^2(n_2)$，且$U, V$相互独立，则称随机变量

$$F = \frac{U/n1}{V/n2}$$

服从第一自由度为n_1，第二自由度为n_2的F分布，记为$F \sim F(n1, n2)$

F分布有如下性质：

1. 若$F \sim F(n1, n2)$，则$1/F \sim F(n2, n1)$

2. 若$X \sim t(n)$，则$X^2 \sim F(1, n)$

3. $F_{1-\alpha}(n1, n2) = \frac{1}{F_\alpha(n2, n1)}$

2.2 四大抽样分布定理

2.2.0.1 （一）

设$X_1, X_2, \ldots, X_n$是来自正态分布总体$N(\mu, \sigma^2)$的简单随机样本，$\bar{X}$是样本均值，S^2是样本方差，则

1. $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

2. $\bar{X}$与S^2相互独立

3. $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

2.2.0.2 （二）

设$X_1, X_2, \ldots, X_{n1}$是来自正态分布总体$N_1(\mu_1, \sigma_1^2)$的简单随机样本,$Y_1, Y_2, \ldots, Y_{n2}$是来自正态分布总体$N_2(\mu_2, \sigma_2^2)$的简单随机样本,样本相互独立$\bar{X}, \bar{Y}$是样本均值，S_1^2, S_2^2是样本方差，则

1. $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n1, n2)$

2. 若$\sigma_1^2 = \sigma_2^2 = \sigma^2$,

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

其中

$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

3 Chapter 7 参数估计

3.1 点估计

3.1.1 （一） 矩法

根据大数定理，当样本容量$n \to +\infty$时，样本矩依概率收敛于相应的总体矩，即

$$A_k \xrightarrow{P} \mu_k, B_k \xrightarrow{P} \nu_k$$

我们将待估参数θ_m写成这些总体矩的函数 [1]

$$\theta_m = h_m(\mu_1, \mu_2, \ldots, \mu_k)$$

根据矩法思想，用样本矩代替总体矩

$$\hat{\theta}_m = h_m(A_1, A_2, \ldots, A_k)$$

将$\hat{\theta}_m$称为参数θ_m的矩估计量

3.1.2 （二）极大似然法

若事件A发生的概率依赖于待估参数θ,且观察到A事件已经发生，则用使事件A发生的概率达到最大的θ值作为θ的估计。

对于离散总体，设其概率分布为$P\{X = x\} = p(x; \theta), \theta \in \Theta, \Theta$是参数空间,则样本$X_1, X_2, \ldots, X_n$取到$x_1, x_2, \ldots, x_n$的概率为$\prod\limits_{i=1}^{n} P\{X_i = x_i\} = \prod\limits_{i=1}^{n} p(x_i; \theta)$

记

$$L(\theta) = \prod_{i=1}^{n} P\{X_i = x_i\} = \prod_{i=1}^{n} p(x_i; \theta)$$

常用微分法

$$l(\theta) = lnL(\theta) = \sum_{i=1}^{n} lnp(x_i; \theta)$$

有

$$\left.\frac{dl(\theta)}{d\theta}\right|_{\theta=\hat{\theta}} = 0$$

3.2 区间估计

设总体为X,$\theta \in \Theta$是待估参数，$X_1, X_2, \ldots, X_n$是来自总体X的样本，统计量$\hat{\theta}_L(X_1, X_2, \ldots, X_n), \hat{\theta}_R(X_1, X_2, \ldots, X_n)$满足$\hat{\theta}_L < \hat{\theta}_R$,且对于给定的$\alpha$，对于任意的$\theta \in \Theta$,有

$$P\{\hat{\theta}_L < \theta < \hat{\theta}_R\} \geq 1 - \alpha$$

则称区间$(\hat{\theta}_L, \hat{\theta}_R)$是参数$\theta$置信水平为$1 - \alpha$的置信区间。

3.2.1 枢轴量

总体X的密度函数为$f(x; \theta)$，其中θ是待估参数，如果样本和参数的函数$G(X_1, X_2, \ldots, X_n, \theta)$的分布完全已知，且不依赖与其他参数，则称$G(X_1, X_2, \ldots, X_n, \theta)$为枢轴量

3.2.2 寻求区间估计的步骤

1. 构造一个枢轴量$G(X_1, X_2, \ldots, X_n, \theta)$

2. 对给定的置信水平$1-\alpha$，根据枢轴量$G(X_1, X_2, \ldots, X_n, \theta)$的分布，选择两常数$a$和$b$,使[2]

$$P_\theta\{a < G(X_1, X_2, \ldots, X_n, \theta) < b\} = 1-\alpha$$

3. 根据分布，取出尽可能小的区间(a, b)，并据此解出$\hat{\theta}_L, \hat{\theta}_R$

习惯上，取a和b满足

$$P_\theta\{G(X_1, X_2, \ldots, X_n, \theta) \le a\} = P_\theta\{G(X_1, X_2, \ldots, X_n, \theta) \ge b\} = \frac{\alpha}{2}$$

3.2.3 枢轴量的选择原则

1. 包含未知参数、未知参数的样本估计、和一个已知参数（有总体参数用总体参数，没有的话用样本量）

2. 分布函数要已知

3. 例子:

   一、单总体 $N\left(\mu, \sigma^2\right)$ 常用枢轴量
   (1) σ^2 已知, 求 μ 的区间估计: $G=\frac{\bar{X}-\mu}{\sigma / \sqrt{n}} \sim N(0,1)$
   (2) σ^2 未知, 求 μ 的区间估计: $\frac{\bar{X}-\mu}{S / \sqrt{n}} \sim t(n-1)$
   (3) μ 未知, 求 σ^2 的区间估计: $\frac{(n-1) S^2}{\sigma^2}=\sum_{i=1}^n\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$
   *(4) μ 已知, 求 σ^2 的区间估计: $\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$

   二、双总体的常用枢轴量

   (1) σ_1^2, σ_2^2已 知, 求 $\mu_1-\mu_2$ 的区间估计: $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$
   (2) $\sigma_1^2=\sigma_2^2$ 未知, 求 $\mu_1-\mu_2$的区间估计:

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim I\left(n_1+n_2-2\right)$$

   (3) μ_1, μ_2 未知, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计: $\frac{S_1^2}{S_2^2} / \frac{\sigma_1^2}{\sigma_2^2} \sim F\left(n_1-1, n_2-1\right)$
   (4) μ_1, μ_2 已知, 求 $\frac{\sigma_1^2}{\sigma_2^2}$ 的区间估计: $\frac{\sum_{m=1}^n (X_1-\mu_1)^2 / n_1}{\sum_{i=1}^2 (Y_1-\mu_2)^2 / n_2} / \frac{\sigma_1^2}{\sigma_2^2} \sim F\left(n_1, n_2\right)$

3.2.4 评价准则

3.2.4.1 （一）无偏性原则

$$E(\hat{\theta}) = \theta,$$

3.2.4.2 （二）有效性原则

若$\hat{\theta}_1, \hat{\theta}_2$都是无偏估计，且$\forall \theta \in \Theta, Var_\theta(\hat{\theta}_1) <= Var_\theta(\hat{\theta}_2)$,要求至少有一个$\theta$使不等号成立，则称$\hat{\theta}_1$比$\hat{\theta}_2$更有效。

3.2.4.3 （三）均方误差原则

要使

$$Mse(\theta, \hat{\theta}) = E[(\hat{\theta}-\theta)^2]$$

尽量小

3.2.4.4 （四）相合性原则

满足

$$\hat{\theta} \xrightarrow{P} \theta$$

1. 有几个待估参数总体矩数量k就用多少 ↩
2. 若对离散随机变量，可能没有办法使概率正好等于$1-\alpha$,因此，a和b应使$P_\theta\{a < G(X_1, X_2, \dots, X_n, \theta) < b\} \geq 1-\alpha$且尽可能接近$1-\alpha$ ↩