概率论与数理统计总结

作为一个成绩呈指数分布的人来写总结，表示压力很大。

王晓峰老师讲课不错，但是很多内容书上没有，也不考。总结主要分如下几个部分：整体框架，重要的或者容易被遗漏的知识点，简单常用的东西就不罗列了，挑了一些典型题目供大家练手，最后是个人的一些复习建议。

整体框架

概率比统计比重大的多，第一章基本上就是从高中到大学的过渡，和实际结合较多，重点是全概率、贝叶斯、条件概率；第二章和第三章是概率也是全书的重点，离散变量——连续变量，分布——密度，期望——方差，都是基础；第四章大数定律就是证明，不会太 BT，知道切比雪夫和马尔科夫两个大数定律就行，中心极限定理只需会应用计算。

统计部分主要是点估计（矩估计、极大似然估计，相合性、无偏性、有效性三个评价标准），区间估计（枢轴量选取），假设检验。

老师说过不考的东西（讲了的）：伽玛、贝塔分布（考特殊情况，见后；个人认为负二项、超几何分布也不会考），弱收敛、强收敛，大样本置信区间、两个正态总体问题，多维正态貌似只考到二维。

TIPS

概率部分

- 从一开始就养成好习惯：区分大小写，大写表示随机变量，小写表示一个值，都用公式书写，不要和高中一样直接写数字，极容易错且错了全没分
- 期望、方差
  - 期望另一种形式：
    $$EX = \int_0^{+\infty} P(X > x)dx - \int_{-\infty}^{0} P(X \le x)dx$$
    主要应用于无密度的连续变量的计算，例题：第二次习题课第 5 题
  - 期望与方差相比最大好处在于其线性性，可以简化计算
  - $\mathrm{Var}(X) = E\big(X - E(X)\big)^2 < E(X-c)^2$，用此方法证明 P89 7,8
- 切比雪夫不等式、马尔科夫不等式的证明，尤其后者证明很典型：构造一个贝努力 0-1 分布，对不等式两边同时取期望，用此方法做 P89 9,10
- 无记忆性（就这两个）：指数分布($min(X_1, X_2, \ldots, X_n) \sim Exp(n\lambda), max(X_1, X_2, \ldots, X_n)$不服从，应用见第三次习题课答案最后一题最后一种解法)、几何分布
- 独立性
  - 证明不独立一般是举出反例（自己带入一组数）
  - 直观判断方法：P154 16 题：X、Y 独立$\Leftrightarrow p(x,y)$可分离变量（边际密度不可为常数）
- X_i独立同分布，最大值$Y = max(X_i)$、最小值$Z = min(X_i)$
  - $p_Y(y) = n[F(y)]^{n-1}p(y)$，$p_Z(z) = n[1-F(z)]^{n-1}p(z)$
  - 离散变量：$P(Y=k) = P(Y<k+1) - P(Y<k)$
    $P(Z=k) = P(Z>k-1) - P(Z>k)$
    例题：第一次习题课第 5 题
  - 对于二变量（均为非负数）常用变换：

PS：这份总结总计耗时 7 个小时，希望大家珍惜。另由于我只复习到概率部分，统计部分尚未整理出典型题目，欢迎提供补充、修正建议。谢谢

$$\text{最大值}U=\frac{|X-Y|}{2}+\frac{X+Y}{2}\text{，最小值}V=-\frac{|X-Y|}{2}+\frac{X+Y}{2}$$

注意到：$X+Y=U+V$，$XY=UV$，取期望后依然成立

例题：第四次习题课第 4、5 题，P182 12、29、44

- 对于(0,1)均匀分布：$E\left(X_{(1)}\right)=\frac{1}{n+1},E\left(X_{(n)}\right)=\frac{n}{n+1}$，其他均匀分布的$E\left(X_{(1)}\right)$和$E\left(X_{(n)}\right)$可通过线性变换求出，十分常用，强烈建议背下来，例题：第五次习题课第 1 题第 3 问
- $X_{(1)}$、$X_{(2)}$、 …、$X_{(n)}$不独立，例题：P300 12

- 伽玛分布$Ga(\alpha,\lambda)$：

$$p(x)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}(x\geq 0)$$

注：此处列出只是为了下面几个式子方便记忆，考试不要求

- 各分布间的特殊关系
  - $Ga(1,\lambda)=Exp(\lambda)$
  - $Ga\left(\frac{n}{2},\frac{1}{2}\right)=\chi^2(n)=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}e^{-\frac{2}{x}}x^{\frac{n}{2}-1}$
  - $X\sim Ga(\alpha,\lambda)\Rightarrow 2\lambda X\sim Ga\left(\alpha,\frac{1}{2}\right)=\chi^2(2\alpha)$
  - $X\sim N(0,1)\Rightarrow X^2\sim\chi^2(1)$ 例题：P165 13
- 可加性（前提：独立）
  - 泊松分布：$X\sim P(\lambda_1),Y\sim P(\lambda_2)\Rightarrow X+Y\sim P(\lambda_1+\lambda_2)$。（注：X-Y 不服从）
  - 二项分布：$X\sim b(n,p),Y\sim b(m,p)\Rightarrow X+Y\sim b(n+m,p)$
  - 正态分布：$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)\Rightarrow X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$
  - 伽玛分布：$X\sim Ga(\alpha_1,\lambda),Y\sim Ga(\alpha_2,\lambda)\Rightarrow X+Y\sim Ga(\alpha_1+\alpha_2,\lambda)$
  - χ^2分布：$\chi^2(n_1)+\chi^2(n_2)+\cdots+\chi^2(n_m)=\chi^2(n_1+n_2+\cdots+n_m)$
  - 指数分布：

$$\sum_{i=1}^{m}Exp(\lambda)=Ga(m,\lambda)$$

- 卷积公式：X、Y 独立，Z = X + Y，

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(z-u)p_Y(u)du$$

- 由联合密度求分布、概率、变换
  - 就是二重积分，强烈建议画出图形区域后积分，十分容易在积分范围上出错，尤其是分段积分，如 P144 7、P164 9、10，另外求边际密度也要注意分段，如 P154 14(1)
  - 变换时注意多对一的情形，此时

$$f_{U,V}(u,v)=\sum f_{X,Y}(x,y)\frac{1}{\left|\frac{\partial(u,v)}{\partial(x,y)}\right|}$$

例题：第三次习题课第 4 题

- 协方差、相关系数
  - 协方差阵应用（估计不考）：P184 39

PS：这份总结总计耗时 7 个小时，希望大家珍惜。另由于我只复习到概率部分，统计部分尚未整理出典型题目，欢迎提供补充、修正建议。谢谢

- 方差、协方差的性质，用于简化计算 P171
- 相关系数是标准化后的协方差，即

$$Cov(X^*,Y^*)=Corr(X,Y)\text{，其中 }X^*=\frac{X-\mu_X}{\sigma_X}\text{，}Y^*=\frac{Y-\mu_Y}{\sigma_Y}$$

- 不等式证明常用：$|Corr(X,Y)|\le 1$，$Var(X)=E(X^2)-EX^2\ge 0$，例题：P185 45

● 二维正态分布
- $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$,联合分布为$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$,$\rho$为相关系数
- 独立等价于不相关
- $aX+bY\sim N(a\mu_1+b\mu_2,a^2\sigma_1^2+b^2\sigma_2^2+2ab\rho\sigma_1\sigma_2)$)（不常用，可无视）

● 重期望公式：$EX=E(E(X|Y))$，例题：第四次习题课第 3 题

● 随机个随机变量和的数学期望（应该不考）：

$$E\left(\sum_{i=1}^{N}X_i\right)=E(X_i)E(N)\text{，例题：P198 13}$$

● 条件期望：例题：P198 12(参考书答案错，见第三次习题课答案最后一题)

● 中心极限定理：随机变量个数 n 应为常数（不能是随机变量），仔细体会第四次习题课第 8 题

统计部分

● 统计和概率区别
- 不区分大小写，小写也可能是随机变量
- 样本中的个体是分布和总体分布相同、相互间独立的随机变量

● 方差（无偏方差）定义：$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$

● 三大分布（均以标准正态为基础）
- χ^2分布：非负值偏态分布，$E\chi^2=n,Var(\chi^2)=2n$，例题：P300 6

  常用结论：$\frac{(n-1)s^2}{\sigma^2}\sim\chi^2(n-1)$，用于构造枢轴量（$\sigma$、$\mu$均未知，估计 μ，用于消去σ）
- F分布：非负值偏态分布
- t分布：与标准正态十分类似

  常用结论：⑴$\frac{\sqrt{n}(\bar{x}-\mu)}{s}\sim t(n-1)$;⑵$t^2\sim F(1,n)$

● 常用结论
- $E\bar{X}=EX$
- $Var\bar{X}=\frac{\sigma^2}{n}$
- $E(s^2)=VarX$

● 点估计
- 矩估计：用样本平均值、方差近似总体均值、方差
  核心式子：

$$\hat{E}(X)=\bar{x}\text{，}\hat{V}ar(X)=s_n^2$$

- 最大似然估计(MLE)：写出似然函数（离散：得到当前样本的概率，连续：密度乘积)，求最大值点（注意被估参数范围）
  - ◆ 有时候不是一个确定点，而是一个范围，例题：P292 9(2)，第五次习题课第 1 题第 2 问

PS：这份总结总计耗时 7 个小时，希望大家珍惜。另由于我只复习到概率部分，统计部分尚未整理出典型题目，欢迎提供补充、修正建议。谢谢

- ◆ 不变性：$g(\theta)$的最大似然估计为$g(\hat{\theta})$
- ■ 评价标准
  - ◆ 相合性：一般点估计都是相合估计

$$\lim_{n\to+\infty} E(\hat{\theta}) = \theta，\quad \lim_{n\to+\infty} \mathrm{Var}(\hat{\theta}) = 0$$

  - ◆ 无偏性：矩估计一般是无偏估计

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

如果不是无偏的，一般方法做线性变换(不影响相合性)修正，例题：第五次习题课第 2 题

  - ◆ 有效性：比方差，小的有效性好，条件：$\hat{\theta}_1$和$\hat{\theta}_2$是无偏估计
  - ◆ 均方误差：$\mathrm{MSE}(\hat{\theta}) = E(\hat{\theta}-\theta)^2$，从这个角度考虑，无偏估计未必比有偏估计好

- ● 区间估计
  - ■ 枢轴量 G 选取
    - ◆ 一定含有样本值（如$\bar{x}, s, x_{(1)}, x_{(n)}$）
    - ◆ 除此外只可含有被估参数，即不可以含有任何其他未知参数
    - ◆ 分布已知
    - ◆ 最好关于被估参数是单调的
  - ■ 适当选择c和d，使$P(c \le G \le d) = 1-\alpha$。尽量使区间长度$d-c$小，但很多情况下做不到，因此取等尾置信区间，即$P(G < c) = P(G > d) = \frac{\alpha}{2}$
  - ■ 根据$c \le G \le d$反解出θ范围
  - ■ 最好用$\bar{x}$（有可加性的分布），有时候$\bar{x}$分布不好求（如均匀分布），则选取$x_{(1)}$或$x_{(n)}$
- ● 假设检验
  - ■ 主观理解：用检测统计量构造一个拒绝域，如果样本观测值落在该区域，则拒绝原假设
  - ■ 和区间估计的关系
    - ◆ 检测统计量和枢轴量选取形式相同
    - ◆ 在拒绝域内认为原假设假；在置信区间里认为估计合理，因此两个区间是互补关系
  - ■ 记弃真概率α(θ)，存伪概率β(θ)，势函数$g(\theta)$，$H_0: \theta \in \Theta_0$，$H_1: \theta \in \Theta_1$

$$g(\theta) = \begin{cases} \alpha(\theta), & \theta \in \Theta_0 \\ 1-\beta(\theta), & \theta \in \Theta_1 \end{cases}$$

两类错误的概率不能都非常小，必会有一个过渡区间，只能通过增大样本容量使图形变陡，减小过渡区间长度。

复习建议

四个资料来源：课本、习题解答、讲义、习题课。我的建议是过一遍书，然后看习题解答上的纲要以及某些人的总结，做一些书上的题目，认认真真把习题课题目和答案全都看一遍到两遍，此时问题基本不大了，可以做一些往年考题，最后有精力再浏览讲义。

最后祝大家好运！

PS：这份总结总计耗时 7 个小时，希望大家珍惜。另由于我只复习到概率部分，统计部分尚未整理出典型题目，欢迎提供补充、修正建议。谢谢