概率论与数理统计复习纲要

整理人：何洪津
杭州电子科技大学 理学院

January 11, 2013

本复习纲要只是对本学期考试重点做了简单的整理，并不一定全面，仅供大家参考。若想学好本课程，还须大家多看书，多思考，多做练习。

第一章：概率论的基本概念

- 事件的运算：通过集合进行运算；
- 概率必须满足如下三条性质:
  1. $0 \le P(A) \le 1$;
  2. $P(S) = 1$, 其中 S 表示整个样本空间；
  3. 若 $A_1, A_2, \cdots, A_k$ 两两互不相容，则 $P(\bigcup_{i=1}^k A_i) = \sum_{i=1}^k P(A_i)$.
- 概率的重要性质：
  1. $P(\emptyset) = 0$;
  2. 若 $A_1, A_2, \cdots, A_n$ 是两两互不相容的事件，则有
  $$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$
  3. 若 $A \subset B$, 则有
  $$P(B - A) = P(B) - P(A), \quad P(B) \ge P(A).$$
  4. 对任意事件 A, 都有 $P(A) \le 1$.

5. 对于任一事件A, 有
$$P(\bar{A}) = 1 - P(A).$$

6. 对 $\forall A, B$，we have
$$P(A \cup B) = P(A) + P(B) - P(AB).$$

- 古典概型的求法；
- **条件概率公式：**
$$P(B|A) = \frac{P(AB)}{P(A)}, \quad P(A) \neq 0.$$

  乘法公式：设 $P(A) > 0$, then we have
$$P(AB) = P(B|A)P(A).$$

  全概率公式：设试验 E 的样本空间为 S, A 为E的事件。$B_1, B_2, \cdots, B_n$ 为 S 的一个划分，且$P(B_i) > 0$, then
$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

  贝叶斯公式：设试验 E 的样本空间为 S, A 为E的事件。$B_1, B_2, \cdots, B_n$ 为 S 的一个划分，且$P(A) > 0$, $P(B_i) > 0$, then
$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}.$$

- **独立：**若 $P(AB) = P(A)P(B)$，则称为 A 与 B 相互独立。注意区别互不相容的定义。会证明当A, B独立时，A 与 $\bar{B}$, $\bar{A}$ 与 B, $\bar{A}$ 与 $\bar{B}$ 都是相互独立的。

第二章：随机变量及其分布

- 离散型随机变量的分布律及其性质，最终要的一条利用 $\sum_{k=1}^{\infty} p_k = 1$ 确定某个参数。
- 三种重要的离散型随机变量的分布律要熟悉：

1. $(0-1)$分布：
$$P(X = k) = p^k(1-p)^{1-k}, \ k = 0, 1, \ (0 < p < 1);$$

2. 二项分布 ($X \sim b(n,p)$)：

$$P\{X=k\} = C_n^k p^k (1-p)^{n-k},\ k=0,1,2,\cdots,n;$$

3. 泊松分布 ($X \sim \pi(\lambda)$)：

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k=0,1,2,\cdots.$$

- 分布函数的定义：$F(x) = P\{X \leq x\}$. 会利用概率密度函数求分布函数（注意分布函数的分段写法，不要漏掉某个部分），即：

$$F(x) = \int_{-\infty}^{x} f(x)dx.$$

  且会利用概率密度函数确定其中包含的未知参数，即利用 $\int_{-\infty}^{\infty} f(x,\theta) = 1$.

- 三种连续型随机变量的分布：

1. 均匀分布（$X \sim U(a,b)$）：概率密度函数为

$$f(x) = \begin{cases} \frac{1}{b-a} & a<x<b, \\ 0 & \text{otherwise.} \end{cases}$$

2. 指数分布概率密度函数($\theta > 0$)：

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & x>0, \\ 0 & \text{otherwise.} \end{cases}$$

3. 正态分布（$X \sim N(\mu,\sigma^2)$）,其概率密度函数为：

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

若 $X \sim N(\mu,\sigma^2)$，则 $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

- 会求简单随机变量的函数的分布，通过分布函数定义求解。

第三章：多维随机变量及其分布

本章内容是第二章内容从一维随机变量到多维随机变量的简单扩充，很多性质，定义基本上可以从第二章直接推广过来。

- 会求边缘分布律；边缘概率密度；条件分布律及条件概率密度；
- 随机变量X和Y的相互独立定义：$P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$ 或者 $f(x,y) = f_X(x)f_Y(y)$.
- 会利用联合概率密度函数确定其中的参数，即利用$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,\theta)dxdy = 1$.
- 会利用联合概率密度函数，求某一点(X,Y) 落在区域G 内的概率, 即

  $$P\{(X,Y) \in G\} = \iint_G f(x,y)dxdy.$$

  做题是要先确定G的区域具体情况，然后确定好积分区域。
- 会求两个随机变量的简单函数的分布。

第四章：随机变量的数字特征

- 会计算数学期望，离散型期望公式

  $$E(X) = \sum_{k=1}^{\infty} x_k p_k,$$

  连续性期望公式：

  $$E(X) = \int_{-\infty}^{\infty} x f(x)dx,$$

  且对于函数$Y = g(X)$的数学期望，直接可由如下公式计算求得：

  $$E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

  数学期望的各条性质要熟悉。

- 会计算方差，

$$D(X)=E\{[X-E(X)]^2\}=E(X^2)-[E(X)]^2=\int_{-\infty}^{\infty}[X-E(X)]^2f(x)dx.$$

熟悉方差的各项性质，并会证明它们。

- 切比雪夫不等式：设$E(X)=\mu$, $D(X)=\sigma^2$, 则对于任意的$\varepsilon>0$, 都有

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

- 协方差

$$\mathrm{Cov}(X,Y)=E\left\{[X-E(X)][Y-E(Y)]\right\}=E(XY)-E(X)E(Y).$$

- 相关系数

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}.$$

相关系数 $|\rho_{XY}|\leq 1$, 且 $|\rho_{XY}|=1$ 的充要条件为，存在 a,b 使得 $P(Y=a+bX)=1$. 当 $\rho_{XY}=0$ 时，称 X,Y不相关。

- 常见分布的均值与方差总结：

Table 1: 常见分布的均值与方差

分布	分布律或密度函数	数学期望	方差
(0-1)分布	$P(X=k)=p^k(1-p)^{1-k},\ k=0,1$	p	$p(1-p)$
二项分布	$P\{X=k\}=C_n^kp^k(1-p)^{n-k},\ k=0,1,\cdots,n$	np	$np(1-p)$
泊松分布	$P(X=k)=\frac{\lambda^ke^{-\lambda}}{k!},\quad k=0,1,\cdots.$	λ	λ
均匀分布$U(a,b)$	$f(x)=\begin{cases}\frac{1}{b-a} & a<x<b,\\ 0 & \text{otherwise.}\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta} & x>0,\\ 0 & \text{otherwise.}\end{cases}$	θ	θ^2
正态分布 $N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty.$	μ	σ^2

第五章：大数定律及中心极限定理

- 独立同分布的中心极限定理：设随机变量$X_1,X_2,\cdots,X_n,\cdots$相互独立，服从同一分布，且具有数学期望和方差： $E(X_k)=\mu$, $D(X_k)=\sigma^2>0$ $(k=1,2,\cdots)$, 则随机变量之和

$\sum_{k=1}^{n} X_k$ 的标准化变量

$$Y_n = \frac{\sum\limits_{k=1}^{n} X_k - E\left(\sum\limits_{k=1}^{n} X_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n} X_k\right)}} = \frac{\sum\limits_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$$

的分布函数$F_n(x)$ 对于任意x满足

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x).$$

会利用此定理求概率问题。在我们的习题集上，有的题目会考虑大于等于零的情况，如综合训练题2第四大题，要求 $P\{X \ge 11\}$, 则可 转化为

$$P\{X \ge 11\} = 1 - P\{X < 11\} = 1 - P\{0 \le X < 11\},$$

在我们的期末考试题中，只考虑 $1 - P\{X < 11\}$ 即可，无需考虑 $X \ge 0$ 的情况。

第六章：样本及抽样分布

- 我们通常说的样本 $X_1, X_2, \cdots, X_n$ 满足两个条件：
  1. 与总体 X 同分布；
  2. $X_1, X_2, \cdots, X_n$ 相互独立。
- 会计算样本 p 分位数，其步骤如下：
  1. 将样本值 $x_1, x_2, \cdots, x_n$ 按从小到大的次序排序成 $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$.
  2. 计算样本 p 分位数，公式为：

$$x_p = \begin{cases} x_{([np]+1)}, & \text{当}np\text{不是整数}, \\ \frac{1}{2}\left[x_{(np)} + x_{(np+1)}\right], & \text{当}np\text{是整数}. \end{cases}$$

- 常用的几个样本统计量：

1. 样本平均值：
$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i.$$
2. 样本方差：
$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right).$$
3. 样本标准差：
$$S=\sqrt{S^2}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}.$$
4. 样本 k 阶（原点）矩
$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k,\quad k=1,2,\cdots.$$
5. 样本 k 阶中心矩
$$B_k=\frac{1}{n}\sum_{i=1}^{n}\left(X_i^k-\overline{X}\right)^k,\quad k=2,3,\cdots.$$

- 三种常用的统计量分布(注意各种分布的假设条件)：

1. χ^2 分布：设 $X_1,X_2,\cdots,X_n$ 是来自总体 $N(0,1)$ 的样本，则称统计量
$$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$$
服从自由度为n的χ^2分布，记为 $\chi^2\sim\chi^2(n)$. 并且，χ^2 分布具有可加性，即对于$\chi_1^2\sim\chi_1^2(n_1)$ 和 $\chi_2^2\sim\chi_2^2(n_2)$, 并且 χ_1^2，χ_2^2 相互独立，则有 $\chi_1^2+\chi_2^2\sim\chi^2(n_1+n_2)$. 另外，若 $\chi^2\sim\chi^2(n)$, 则有 $E(\chi^2)=n,\ D(\chi^2)=2n.$
2. t 分布：设 $X\sim N(0,1)$, $Y\sim\chi^2(n)$, 且X,Y 相互独立，则称随机变量
$$t=\frac{X}{\sqrt{Y/n}}$$
服从自由度为 n 的 t 分布，记为 $t\sim t(n)$. t 分布的概率密度函数具有对称性。
3. 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
$$F=\frac{U/n_1}{V/n_2}$$

服从自由度为 (n_1, n_2) 的 F 分布，记为 $F \sim F(n_1, n_2)$.

- 正态总体的样本均值与样本方差的分布：

1. 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$ 是样本均值，则有

$$\overline{X} \sim N(\mu, \sigma^2/n).$$

2. 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$ 和 S^2 分别是样本均值和样本方差，则有

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1); \quad 且 \quad \overline{X} 与 S^2 \ 相互独立.$$

3. 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$ 和 S^2 分别是样本均值和样本方差，则有

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

第七章：参数估计

- 矩估计法其基本思想是：在总体存在所需要的各阶矩条件下，用样本的各阶矩估计总体的相应的各阶矩。由于总体的分布类型已知， 总体的各阶矩可表示成未知参数的已知函数，这样样本的各阶矩就与未知参数联系起来了，从而可得到未知参数的估计。（有多少个 未知的参数，我们就使用多少阶矩）。

- 最大似然估计其基本思想是：根据样本的具体情况，选择参数θ的估计$\hat{\theta}$使得该样本发生的概率最大， 即求解如下似然函数的最大值所对应的参数θ:

$$(离散型) \quad L(x_1, x_2, \cdots, x_n; \hat{\theta}) = \max_{\theta \in \Theta} L(x_1, x_2, \cdots, x_n; \theta) = \max_{\theta \in \Theta} \prod_{i=1}^{n} p(x_i; \theta);$$

或

$$(连续型) \quad L(x_1, x_2, \cdots, x_n; \hat{\theta}) = \max_{\theta \in \Theta} L(x_1, x_2, \cdots, x_n; \theta) = \max_{\theta \in \Theta} \prod_{i=1}^{n} f(x_i; \theta).$$

要求$\hat{\theta}$使得似然函数取最大值，只需对似然函数 $L(x;\theta)$ 或者似然函数的对数函数 $\ln L(x;\theta)$ 关于 θ 求偏导令其等于零，解出 θ 即可，当然碰到不可导的情况，我们只能通过定义进行求解了。

- 若估计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ 的数学期望 $E(\hat{\theta})$ 存在， 且对于任意的 $\theta \in \Theta$ 有

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

则称 $\hat{\theta}$ 是 θ 的无偏估计量.

- 区间估计公式总结：

Table 2: 单个正态总体均值、方差的置信区间与单侧置信限(置信水平为$1-\alpha$)

待估参数	其他参数	枢轴量的分布	置信区间	单侧置信限
μ	σ^2 已知	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(\overline{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$	$\bar{\mu} = \overline{X} + \frac{\sigma}{\sqrt{n}} z_\alpha$, $\underline{\mu} = \overline{X} - \frac{\sigma}{\sqrt{n}} z_\alpha$
μ	σ^2 未知	$t = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\overline{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right)$	$\bar{\mu} = \overline{X} + \frac{S}{\sqrt{n}} t_\alpha(n-1)$, $\underline{\mu} = \overline{X} - \frac{S}{\sqrt{n}} t_\alpha(n-1)$
σ^2	μ 未知	$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$	$\bar{\sigma}^2 = \frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$, $\underline{\sigma}^2 = \frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}$

第八章：假设检验

- 两类错误：第一类错误也叫弃真错误，即当原假设H_0为真时，我们却拒绝了假设H_0，我们也通常只考虑这类错误。 若设定显著性水平为 α, 则第一类错误可表示为

$$P(\text{拒绝}H_0|H_0\text{为真}) = \alpha.$$

第二类错误为受伪错误（亦叫取伪错误），即当原假设H_0为假时，我们却接受了假设H_0。

- 假设检验可分为双边假设检验和单边假设检验，当题目出现是否发生显著变化时，实际上是双边假设检验问题，此时我们假设

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0;$$

若题目出现了是否偏大、偏小等字眼，说明是单边假设问题，此时我们针对是否 偏大的问题，要假设

$$H_0 : \mu \leq \mu_0, \quad H_1 : \mu > \mu_0;$$

若问是否偏小的话，则要假设

$$H_0 : \mu \geq \mu_0, \quad H_1 : \mu < \mu_0.$$

- 正态总体均值、方差的检验方法：

Table 3: 正态总体均值、方差的检验方法（显著性水平为α）

原假设 H_0	检验统计量	备择假设 H_1	拒绝域
$\mu \leq \mu_0$ $\mu \geq \mu_0$ $\mu = \mu_0$ (σ^2 已知)	$Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$z \geq z_\alpha$ $z \leq -z_\alpha$ $\|z\| \geq z_{\alpha/2}$
$\mu \leq \mu_0$ $\mu \geq \mu_0$ $\mu = \mu_0$ (σ^2 未知)	$t = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$t \geq t_\alpha(n-1)$ $t \leq -t_\alpha(n-1)$ $\|t\| \geq t_{\alpha/2}(n-1)$
$\sigma^2 \leq \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$ $\sigma^2 = \sigma_0^2$ (μ 未知)	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 \geq \chi_\alpha^2(n-1)$ $\chi^2 \leq \chi_{1-\alpha}^2(n-1)$ $\chi^2 \geq \chi_{\alpha/2}^2(n-1)$ 或 $\chi^2 \leq \chi_{1-\alpha/2}^2(n-1)$