概率统计期末复习

冬季期末考试

（1）线上

（2）开卷，独立完成

（3）50 道选择判断题

（4）平时:期末 = 3:7

（5）25 日或 26 日，具体时间等通知

第一章　概率论的基本概念

§1 随机试验 §2 样本空间、随机试验

随机试验

样本空间、事件

子事件、和事件、积事件、差事件

必然事件、不可能事件、互斥（互不相容）、对立

运算律　De Morgan 律等

事件的表示与化简

§3 频率与概率

频率、概率的统计定义

概率的公理化定义　非负性、规范性、可列可加性

概率的性质　$P(\varnothing)=0$　有限可加性

$$P(A-B)=P(A)-P(AB)$$

$B \subset A$ 时,　$P(A-B)=P(A)-P(B)$　$P(A)\le P(B)$

$$P(\bar{A})=1-P(A)$$　$P(A)\le 1$

加法公式　$P(A\cup B)=P(A)+P(B)-P(AB)$　推广

§4 古典概型

古典概型定义

$$P(A)=\frac{A\text{ 包含的样本点数}}{\text{样本空间中的样本点总数}}$$

计数法则　　乘法法则　　加法法则

排列　　全排列

组合　　分组组合

有放回抽样　无放回抽样

§5 条件概率

条件概率 $P(A\mid B):=\dfrac{P(AB)}{P(B)}$ $\boxed{P(B)>0}$

条件概率的性质 非负性、规范性、可列可加性

乘法公式

$P(AB)=P(B)P(A\mid B)$ $\boxed{P(B)>0}$

$P(AB)=P(A)P(B\mid A)$ $\boxed{P(A)>0}$

推广 $P(ABC)=P(C\mid AB)P(B\mid A)P(A)$

全概率公式 $P(A)=\sum_{i=1}^{n}P(B_i)P(A\mid B_i)$ 完备事件组

贝叶斯公式 $P(B_j\mid A)=\dfrac{P(AB_j)}{P(A)}=\dfrac{P(B_j)P(A\mid B_j)}{\sum_{i=1}^{n}P(B_i)P(A\mid B_i)}$

§6 独立性

独立性定义 $P(AB) = P(A)P(B)$

$\Longleftrightarrow$ $P(A \mid B) = P(A)$ $\boxed{P(B) > 0}$

$\Longleftrightarrow$ $P(B \mid A) = P(B)$ $\boxed{P(A) > 0}$

$\boxed{P(A) > 0,\ P(B) > 0}$

独立 $\Longrightarrow$ 不互斥

互斥 $\Longrightarrow$ 不独立

多个事件的独立性
$$
\begin{cases}
P(AB) = P(A)P(B) \\
P(BC) = P(B)P(C) \\
P(CA) = P(C)P(A) \\
P(ABC) = P(A)P(B)P(C)
\end{cases}
$$

第二章　随机变量及其分布

§1 随机变量

样本空间上的实值单值函数

离散型　　连续型　　非离散非连续型

把事件用随机变量的关系式表达出来

§2 离散型随机变量及其分布律

分布律　所有取值 ✚ 每个取值的概率

列表法　图示法　公式法

性质　(1) $p_k \geq 0$,　(2) $\sum p_k = 1$

(0-1) 分布　$X \sim b(1, p)$　$0 \leq p \leq 1$

$$P(X = k) = p^k (1-p)^{1-k},\ k = 0,1$$

X	0	1
$P(X)$	$1-p$	p

$EX = p$　　$DX = p(1-p)$

伯努利试验 (1) 只考虑成功失败 (2) 可重复性 (3) 独立性

二项分布

$$X \sim b(n,p) \qquad 0 \le p \le 1$$

$$P(X=k) = \mathrm{C}_n^k\, p^k \left(1-p\right)^{n-k},\ k=0,1,\cdots,n$$

$$EX = np \qquad DX = np(1-p)$$

泊松分布 (1) 平稳性 (2) 无后效性 (3) 普通性

相同长度等概率 不重叠独立 充分小只发生一次

$$X \sim \pi(\lambda) \qquad \lambda > 0$$

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

$$EX = DX = \lambda$$

§3 随机变量的分布函数

分布函数 $F(x) := P(X \le x)$

性质 (1) 不减函数

(2) $F(-\infty) = 0,\ F(+\infty) = 1$

(3) 右连续

$$P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$$

§4 连续型随机变量及其概率密度

概率密度　描述每一点附近概率的函数 $f(x)$

$$F(x)=\int_{-\infty}^{x} f(t)\,dt = P(X\le x)$$

性质　(1) $f(x)\ge 0$,　(2) $\int_{-\infty}^{+\infty} f(x)\,dx=1$

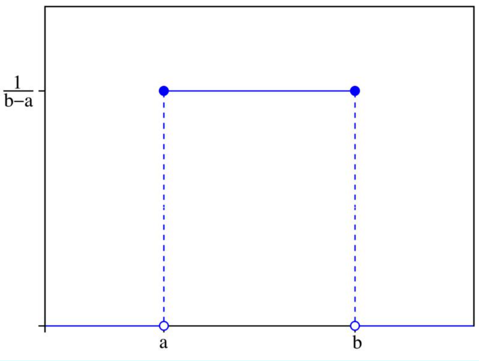
(3) $P(a<X\le b)=\int_{a}^{b} f(x)\,dx=F(b)-F(a)$

(4) $F'(x)=f(x)$　$f(x)$ 在点 x 连续

$P(x=a)=0$　$P(A)=0 \not\Rightarrow A=\varnothing$　$P(A)=1 \not\Rightarrow A=S$

$$P(a\le X\le b)=P(a<X\le b)=P(a\le X<b)=P(a<X<b)$$

均匀分布

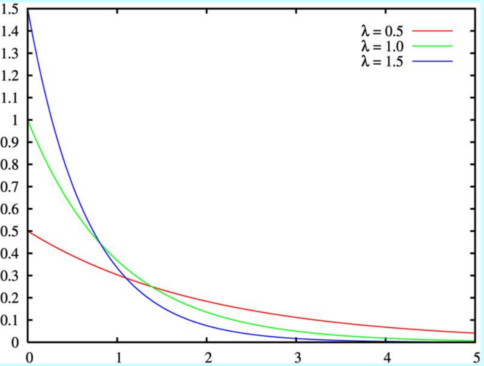


$$X \sim U(a,b)$$

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其它}\end{cases} \qquad F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\le x<b\\ 1, & b\le x\end{cases}$$

$$EX=\frac{a+b}{2} \qquad DX=\frac{(b-a)^2}{12}$$

指数分布



$$X \sim e(\theta) \qquad \theta>0$$

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0\\ 0, & x\le 0\end{cases} \qquad F(x)=\begin{cases}1-e^{-\frac{x}{\theta}}, & x>0\\ 0, & x\le 0\end{cases}$$

$$EX=\theta \qquad DX=\theta^2$$

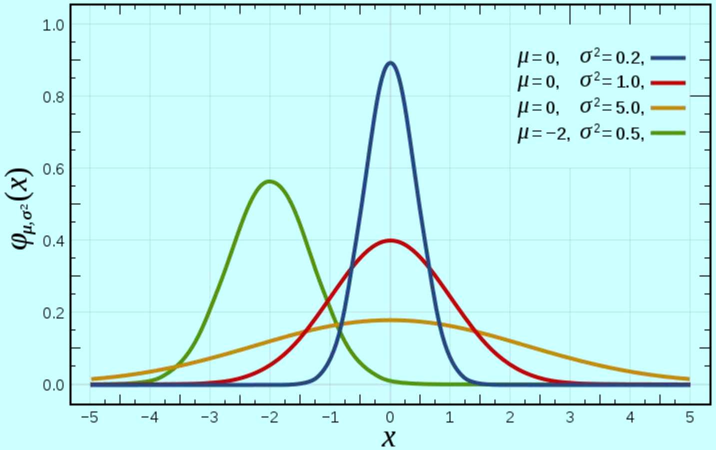
正态分布

$$X \sim N(\mu,\sigma^2) \quad \sigma > 0$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$EX = \mu \qquad DX = \sigma^2$$



关于 μ 对称

μ 处最大值, 左增右减

$\mu \pm \sigma$ 拐点

x 轴为渐近线

标准正态分布

$$X \sim N(0,1)$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,dt$$

(1) $\Phi(0) = 0.5$

(2) $\Phi(-x) = 1 - \Phi(x)$

$X \sim N(\mu,\sigma^2)$ 时,

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

3σ 准则

§5 随机变量的函数的分布

离散型　　合并

连续型　　(1) $F_Y(y)=P(Y\le y)=P(g(X)\le y)$　　先求分布函数

$$(2)\ f_Y(y)=\frac{dF_Y(y)}{dy}$$

再求导得密度

$$f_Y(y)=\begin{cases} f[h(y)]\left|\dfrac{dh(y)}{dy}\right|, & \alpha<y<\beta \\ 0, & \text{其它} \end{cases}$$

$y=g(x)$ 可导

$x=h(y)$ 是 $y=g(x)$ 的反函数

X 的分布函数 $F(x)$ 严格单调连续 $\Longrightarrow$ $Y=F(x)\sim U(0,1)$

第三章　多维随机变量及其分布

§1 二维随机变量

分布函数　$F(x,y) := P\big((X \le x) \cap (Y \le y)\big) =: P(X \le x, Y \le y)$

$$P\left(x_1 < X \le x_2, y_1 < Y \le y_2\right)$$
$$= F\left(x_2, y_2\right) - F\left(x_2, y_1\right) - F\left(x_1, y_2\right) + F\left(x_1, y_1\right)$$

性质　(1) 对 x (或对 y)不减函数

(2) $F(x,-\infty)=0,\ F(-\infty,y)=0$

$F(-\infty,-\infty)=0,\ F(+\infty,+\infty)=1$

(3) 对 x (或对 y) 右连续

二维离散型 $P(X=x_i, Y=y_j)=p_{ij}, \quad i,j=1,2,\ldots$

联合分布律 列表法 公式法

性质 (1) $p_{ij}\geq 0$, (2) $\sum_i \sum_j p_{ij}=1$

二维连续型 $F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\, dudv$

性质 (1) $f(x,y)\geq 0$, (2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\, dxdy=1$

(3) $P\left((X,Y)\in G\right)=\iint_G f\left(x,y\right)dxdy$

(4) $f(x,y)=\dfrac{\partial^2 F(x,y)}{\partial x \partial y}$ $f(x,y)$ 在点 (x,y) 连续

§2 边缘分布

边缘分布函数

$$F_X(x) = P(X \le x) = P(X \le x, Y < +\infty) = F(x, +\infty)$$

$$F_Y(y) = P(Y \le y) = P(X \le +\infty, Y < y) = F(+\infty, y)$$

离散型 $P(X = x_i) = \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{j=1}^{\infty} p_{ij} =: p_{i\bullet}$

$$P(Y = y_j) = \sum_{i=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{i=1}^{\infty} p_{ij} =: p_{\bullet j}$$

连续型 $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

二维均匀分布

二维正态分布 $(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

边缘分布

$$X \sim N(\mu_1,\sigma_1^2)$$

$$Y \sim N(\mu_2,\sigma_2^2)$$

相关系数 ρ

由联合分布可以确定边缘分布

由边缘分布一般不能确定联合分布

§3 条件分布

条件分布律

$$P(X=x_i \mid Y=y_j) = \frac{P\left(X=x_i, Y=y_j\right)}{P\left(Y=y_j\right)} = \frac{p_{ij}}{p_{\bullet j}}$$

$$P(Y=y_j \mid X=x_i) = \frac{P\left(X=x_i, Y=y_j\right)}{P\left(X=x_i\right)} = \frac{p_{ij}}{p_{i\bullet}}$$

条件密度

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} \qquad F_{X|Y}\left(x \middle| y\right) = \int_{-\infty}^{x} \frac{f\left(x,y\right)}{f_Y\left(y\right)} dx$$

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)} \qquad F_{Y|X}\left(y \middle| x\right) = \int_{-\infty}^{y} \frac{f\left(x,y\right)}{f_X\left(x\right)} dy$$

§4 相互独立的随机变量

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$$

分布函数表示 $F(x, y) = F_X(x)F_Y(y)$

离散型表示 $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$

$$p_{ij} = p_{i\bullet}p_{\bullet j}$$

连续型表示 $f(x, y) = f_X(x)f_Y(y)$ 几乎处处成立

有放回抽样 $\Rightarrow$ 独立

无放回抽样 $\Rightarrow$ 不独立

§5 两个随机变量的函数的分布

$Z = X + Y$

离散型 $$P(Z=r)=\sum_{i=0}^{r}P(X=i,Y=r-i)$$

连续型

卷积公式

$$\begin{cases} f_Z(z)=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)\,dy \\ f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\,dx \end{cases}$$

独立

$M=\max\{X,Y\}$ $$F_M(z)=F_X(z)F_Y(z)$$ 独立

$N=\min\{X,Y\}$ $$F_N(z)=1-[1-F_X(z)][1-F_Y(z)]$$ 独立

$X_1,\ldots,X_n \sim b(1,p)$ 独立同分布 $\Longrightarrow$ $\sum_{i=1}^{n} X_i \sim b(n,p)$

$X \sim b(m,p),\ Y \sim b(n,p)$ 独立 $\Longrightarrow$ $X+Y \sim b(m+n,p)$

$X \sim \pi(\lambda_1),\quad Y \sim \pi(\lambda_2)$ 独立 $\Longrightarrow$ $X+Y \sim \pi(\lambda_1+\lambda_2)$

$X \sim N(\mu_1,\sigma_1^2),\ Y \sim N(\mu_2,\sigma_2^2)$ 独立 $\Longrightarrow$

$$aX+bY \sim N(a\mu_1+b\mu_2, a^2\sigma_1^2+b^2\sigma_2^2)$$

$X \sim e(\theta_1),\ Y \sim e(\theta_2)$ 独立 $\Longrightarrow$ $\min\{X,Y\} \sim e\left(\dfrac{\theta_1\theta_2}{\theta_1+\theta_2}\right)$

第4章 随机变量的数字特征

（一）要求:

(1)掌握数学期望、方差的定义及其计算;

(2)掌握数学期望、方差的公式及其证明;

（二）数学期望 口诀：取值 乘以 取这个值的概率 之和。

$$EX=\sum_{i=1}^{\infty}x_i p_i \qquad EX=\int_{-\infty}^{+\infty}xf(x)dx$$

$$EX=\sum_{i=1}^{\infty}g(x_i)p_i \qquad EX=\int_{-\infty}^{+\infty}g(x)f(x)dx$$

$$Eg(X,Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}g(x_i,y_j)p_{ij}$$

$$Eg(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy=\iint_{R^2}g(x,y)f(x,y)dxdy$$

六、数学期望的性质

$$\left.\begin{array}{ll}\square\ E(c)=c & \text{——常数} \\ \square\ E(cX)=cE(X) & \\ \square\ E(X+Y)=E(X)+E(Y) & \end{array}\right\}\Rightarrow \text{线性性质}$$

$$E\left(\sum_{i=1}^{n} c_i X_i + C\right) = \sum_{i=1}^{n} c_i E(X_i) + C$$

- 当X,Y独立时，$E(XY)=E(X)E(Y)$.

反之未必成立，即

若$E(XY)=E(X)E(Y)$，X,Y不一定独立

（四）方差

$$D(X)=E\left[X-E(X)\right]^2$$

离差（偏差）平方的数学期望

$\sqrt{DX}$ 称为X的均方差或标准差。

（五）方差性质

$$D(X)=E(X^2)-[E(X)]^2$$

1、$D(c)=0$

2、$D(cX)=c^2D(X)$

$$D(c_1X+c_2)=c_1{}^2D(X)$$

3、 $D(X+Y)=DX+DY+2Cov(X,Y)$

特别地，若X,Y相互独立，则

$$D(C_1X\pm C_2Y)=C_1^2D(X)+C_2^2D(Y)$$

第4章 几类重要的分布 6-10分

(1) 0-1分布（两点分布）

$$P\{X=0\}=1-p=q, \quad P\{X=1\}=p$$

或 $p(x)=P\{X=x\}=p^x q^{1-x},\ x=0,\ 1$

或

X	0	1
p_k	$1-p$	p

记作 $X \sim B(1,\ p)$ ，或 X ~‘0-1分布’

(2) 二项分布

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

记作 $X \sim B(n,\ p)$

(3) 泊松分布(Poisson)

若X的分布为 $P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda},\ k=0,\ 1,\ 2,\ \dots\ ,$

其中$\lambda>0$是常数, 则称X服从参数为λ的泊松分布. 记为$X\sim\pi(\lambda)$或$P(\lambda)$.

(4) 几何分布

$$P\{X=k\}=q^{k-1}p,\quad k=1,\ 2,\ \dots$$

(5) 超几何分布

设一批产品有N个，其中有M个次品。从这批产品中任意抽取n个，则取出的产品中次品数X的分布律为

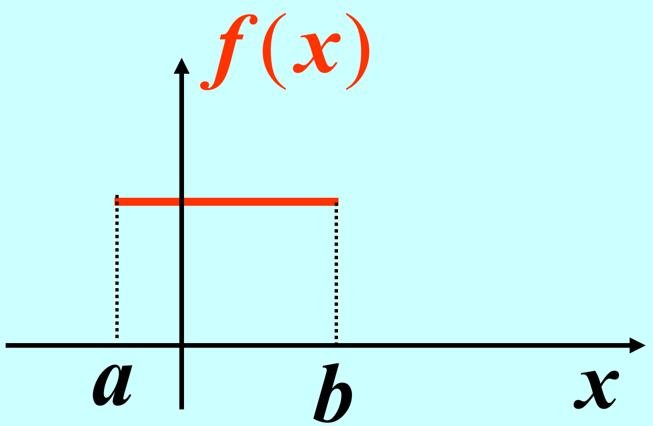
$$P(X=k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n},$$

$$\max(0,\ n-N+M) \le k \le \min(n,\ M)$$

称为X服从参数为p的超几何分布. 记为 $X \sim H(n, M, N)$

(6) 均匀分布：

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



则称X服从区间(a, b)上的均匀分布，记作：$X \sim U(a, b)$

（7）指数分布

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & \text{其它}\end{cases}$$ 其中**λ > 0** 为常数

记为 $X \sim E(\lambda)$.

(8)正态分布 或 高斯分布

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

记为 $X \sim N(\mu,\sigma^2)$.

（三）常见随机变量 的数学期望

分布	概率分布列	期望	方差
参数为p 的 0-1分布 $B(1, p)$	$P(X=1)=p$ $P(X=0)=1-p$	p	pq
二项分布 $B(n,p)$	$P(X=k)=C_n^k p^k (1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np	npq
泊松分布 $P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\cdots$	λ	λ

（三）常见随机变量 的数学期望

分布	密度	期望	方差
区间(a,b)上的均匀分布 U(a, b)	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & 其它\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
指数分布 $E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & 其它\end{cases}$	λ^{-1}	λ^{-2}
正态分布 $N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

第5章 中心极限定理

（一）要求:

(1)熟记依概率收敛的定义;
(2)理解什么是大数定律，什么是中心极限定理?
(3)会写切比雪夫不等式、大数定律。

（二）依概率收敛:

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P\{|Y_n - a| \geq \varepsilon|\} = 0$$

则称序列$Y_1, Y_2, Y_n, \ldots$依概率收敛于a,

记为 $$Y_n \xrightarrow{P} a.$$

（三）切比雪夫不等式（要求会写，会证）：

设随机变量X 的期望EX及方差DX存在，则对任意的 ε>0，有

$$P(|X-EX|\geq\varepsilon)\leq\frac{DX}{\varepsilon^2}$$

偏差较大
的概率较小

证明：（1）设X 为连续型随机变量，其密度函数为f(x)

$$P(|X-EX|\geq\varepsilon)=\int_{|x-EX|\geq\varepsilon}f(x)dx$$

扩大被积函数

$$\leq\int_{|x-EX|\geq\varepsilon}\frac{(x-EX)^2}{\varepsilon^2}f(x)dx$$

$$\leq\int_{-\infty}^{+\infty}\frac{(x-EX)^2}{\varepsilon^2}f(x)dx=\frac{DX}{\varepsilon^2}$$

扩大积分限

四、契比雪夫(大数定律)定理 （要求会证）

设$\{X_n\}$为相互独立的随机变量序列，$E(X_i)$存在，$D(X_i)\leqslant M$ ，（i=1，2，…）则对 ∀ε>0，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right\} = 1$$

证明：（用契比雪夫不等式）

（五）要熟记几个大数定律，中心极限定理:

熟记名称及内容:

契比雪夫(大数定律)定理，辛钦大数定律，伯努利大数定律

熟记名称及内容的中心极限定理:

林德伯格-列维（独立同分布），

棣莫佛-拉普拉斯。

第6章 样本及抽样分布

（一）要求:

(1)理解把握总体、个体、样本、统计量的定义

(2)熟悉常用统计量：
样本均值、方差、一、二阶原点矩中心矩

(3)三大分布的定义、图形、查表

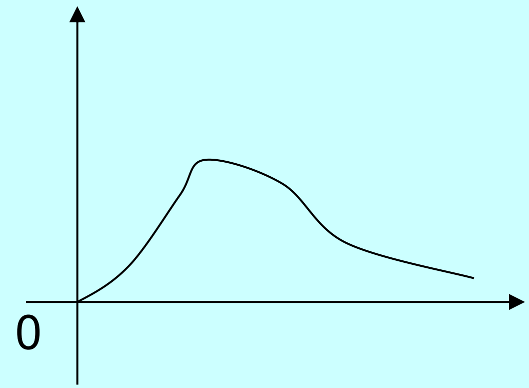
(4)熟练掌握抽样分布定理

（二）三大分布:

定义1（卡方分布）：

$$\chi^2 = X_1^{\ 2} + X_2^{\ 2} + \cdots + X_n^{\ 2} \sim \chi^2(n)$$

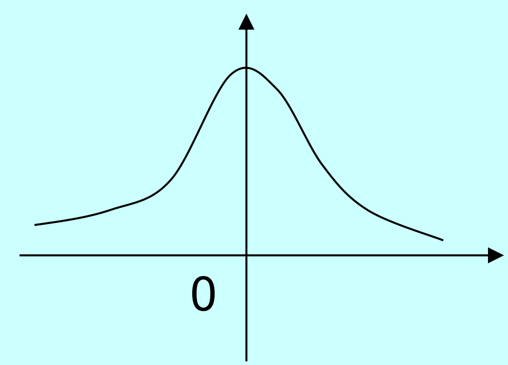
其中$X_i \sim N(0,1)$



定义2（t-分布）：

$$T = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

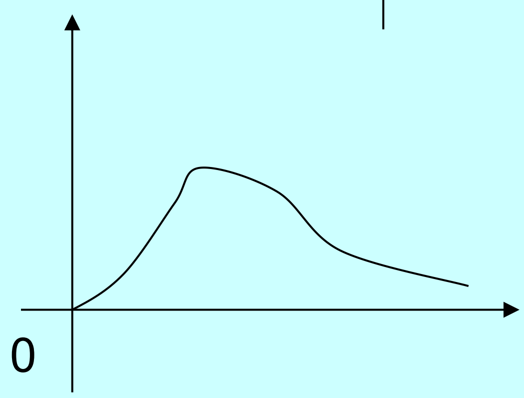
其中 $X \sim N(0,1),\ Y \sim \chi^2(n)$



定义3（F-分布）：

$$F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$$

其中 $X \sim \chi^2(n_1),\ Y \sim \chi^2(n_2)$



（三）抽样分布定理:

总体$X \sim N(\mu, \sigma^2)$，则

$$Th1: U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,\ 1)$$

$$Th4: T=\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

$$\text{Th2}: \quad \chi_1^2=\frac{n\tilde{S}^2}{\sigma^2} \sim \chi^2(n)$$

$$\text{Th3}: \quad \chi_2^2=\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

其中 $\tilde{S}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$

$$Th7: \quad F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}=\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \sim F(n_1-1, n_2-1)$$

要求熟记这5个定理

第7章 参数估计

（一）要求:

(1)掌握点估计中的矩估计及 极大似然估计（对离散及连续）；

(2)掌握无偏性，有效性，一致性的定义，及应用;

(3)区间估计公式（简单的不复杂的）。

（二）矩估计: 用样本矩去估总体矩。

矩估计法的具体步骤:

0、确定分布$F(x;\theta_1,\theta_2,\ldots,\theta_m)$（如果未知）；

1、计算总体的k阶原点矩EX^k；（至多到2阶）

2、用 $M_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k=EX^k$ 建立方程组;

3、解方程组；4、戴帽； 5、回答（当需要时）。

例 设总体X在区间$[0, \theta]$上服从均匀分布，其中$\theta>0$ 为未知参数，若取样本为$X_1,\cdots, X_n$，试求θ的矩估计.

解：

0、确定分布$F(x;\theta_1, \theta_2, \ldots, \theta_m)$（如果未知）；

1、计算总体的k阶原点矩EX^k

$$f(x)=\begin{cases}\frac{1}{\theta}, & 0\le x\le\theta\\ 0, & \text{其他}\end{cases}$$

$$EX=\int_{-\infty}^{+\infty}xf(x)dx$$

$$=\int_0^{\theta}\frac{x}{\theta}dx=\left.\frac{x^2}{2\theta}\right|_0^{\theta}=\frac{\theta}{2},$$

$$EX=\frac{0+\theta}{2}=\frac{\theta}{2},$$

2、用 $\boxed{M_k=\frac{1}{n}\sum_{i=1}^{n}X_i^{\,k}=EX^k}$ **建立方程组；**

$$M_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}=EX=\frac{\theta}{2}$$

3、解方程组； **4、戴帽；** **5、回答（当需要时）；**

$$\Rightarrow\hat{\theta}=2\overline{X}$$ 为所求θ的矩估计(量)。

（三）极大似然估计:

极大似然原理：
总体参数应有利于样本观察值的出现。

求最大似然估计量的步骤:

0、确定分布（若未知）；

1、写出**似然函数并化简**

$$L(\theta)=L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}p(x_i;\theta)\ ;\qquad L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)\ ;$$

2、写出对数**似然函数并化简**

$$\ln L(\theta)=\sum_{i=1}^{n}\ln p(x_i;\theta)\quad ;\qquad \ln L(\theta)=\sum_{i=1}^{n}\ln f(x_i;\theta)\quad ;$$

3、 对 θ 求导 $\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}$，并令 $\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=0$ ；

4、解方程（组），求驻点，就得 θ 的极大似然估计。

（四）三种评价标准:

(1) 无偏性

$E(\hat{\theta}) = \theta$ 则称$\hat{\theta}$ 为θ 的无偏估计.

(2) 有效性

若 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$

则称 $\hat{\theta}_1$ 较（比）$\hat{\theta}_2$ 有效。

(3) 一致性

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

则称$\hat{\theta}_n$是总体参数θ的一致(或相合)估计量

（五）区间估计:

定义7.4.1： 设X是以θ为未知参数的总体， X_1，X_2，…，X_n来自总体的样本。如果对于小概率α（一般取**0.1, 0.05**等），存在两个统计量 $T_1(X_1, X_2, \ldots, X_n)$， $T_2(X_1, X_2, \ldots, X_n)$，使得

$$P\{T_1 \le \theta \le T_2\} = 1-\alpha,$$

则称随机区间 $[T_1, T_2]$为θ为的区间估计, T_1为置信下限, T_2为置信上限， $1-\alpha$置信水平（或置信度）；$[T_1, T_2]$又称为参数θ的置信区间。

求置信区间的一般步骤如下（枢轴统计量法）：

1. 明确问题, 是对什么参数求置信区间?

   置信水平 $1-\alpha$ 是多少?

2. 寻找参数 θ 的一个良好的统计量（或随机变量，称为枢轴统计量）

$$T(\theta) = T(X_1, X_2, \ldots X_n;\ \theta)$$

（1）含待估参数

（2）不含其他未知参数

（3）分布为已知（可查表）.

3. 对于给定的置信水平1-α，根据T的分布，确定常数a, b，使得 $P(a \le T(\theta) \le b) = 1-\alpha$

4. 对 “$a \le T(\theta) \le b$” 作等价变形, 得到如下形式:

$$\hat{T}_1 \le \theta \le \hat{T}_2$$

从而随机区间 $[\ T_1\ ,\ T_2\ ]$就是θ的置信水平为1- α的置信区间。

（七）置信区间公式:

即 μ 的双侧置信区间为

$\sigma=\sigma_0$已知 $$[\bar{X}-\frac{\sigma_0}{\sqrt{n}}u_{1-\alpha/2}, \bar{X}+\frac{\sigma_0}{\sqrt{n}}u_{1-\alpha/2}]$$

σ未知 $$[\bar{X}-\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1), \ \bar{X}+\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)]$$

从而得σ^2的置信水平

为1-α的（双侧）置信区间为

μ未知 $$\left[\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}\right]$$

第9章 假设检验

（一）要求:

(1)理解小概率原理;

(2)熟悉假设检验步骤（并能熟练应用）;

(3)理解两类误差。

（二）假设检验步骤:

1、根据题意，提出恰当的假设

（1）对哪个参数在什么条件下进行检验

（2）单侧，还是双侧检验;

备择假设的方向与提问方向一致！

2、确定检验统计量。

3、确定拒绝域: **拒绝域方向**与备择假设的方向一致。

4、计算并判断;

5、回答（解释）。

（1）对μ进行检验

是否有显著差异

是否明显减小

是否明显增大

假设	双侧检验	左侧检验	右侧检验
假设形式	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$
统计量	σ已知：	$U = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$	
	σ未知：	$T = \dfrac{\overline{X} - \mu_0}{S / \sqrt{n}}$	
拒绝域	$\lvert U \rvert > u_{1-\alpha/2}$	$U < -u_{1-\alpha}$	$U > u_{1-\alpha}$
	$\lvert T \rvert > t_{1-\alpha/2}(n-1)$	$T < -t_{1-\alpha}(n-1)$	$T > t_{1-\alpha}(n-1)$

（2）对σ^2 进行检验

是否有显著差异

是否明显减小

是否明显增大

假设	双侧检验	左侧检验	右侧检验
假设形式	$H_0:\ \sigma^2=\sigma^2_0$ $H_1:\ \sigma^2\neq\sigma^2_0$	$H_0:\ \sigma^2\geq\sigma^2_0$ $H_1:\ \sigma^2<\sigma^2_0$	$H_0:\ \sigma^2\leq\sigma^2_0$ $H_1:\ \sigma^2>\sigma^2_0$
统计量	μ 已知：	$\chi^2=\dfrac{n\widetilde{S}^2}{\sigma^2}=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$	
	μ未知：	$\chi^2=\dfrac{(n-1)S^2}{\sigma^2}=\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}\sim\chi^2(n-1)$	
拒绝域	$\chi^2<\chi^2_{\alpha/2}$或$\chi^2>\chi^2_{1-\alpha/2}$	$\chi^2<\chi^2_{\alpha}$	$\chi^2>\chi^2_{1-\alpha}$

（3）均值差方差比的检验自己总结

是否有显著差异

是否明显减小

是否明显增大减小

假设	双侧检验	左侧检验	右侧检验
假设形式	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$
统计量	已知：		
	未知：		
拒绝域			

（三） 两类错误

第一类错误（又称为弃真错误）

原假设本身为真，但由于小概率事件发生，错误的拒绝了原假设而犯的错误。

第二类错误（又称为存伪错误）

原假设本身为假，但却错误地接受了原假设而犯的错误。

在样本容量一定的情况下，两类错误不能同时减小。

即犯第一类错误的概率减小必然导致犯第二类错误的概率增加，反之亦然！