概率统计各章节总结

第一章

概率的计算

1 ）统计定义：$f_n(A) \xrightarrow[n\to\infty]{} 稳定值 = P(A)$

2 ）概率的性质： 1~5

3 ）等可能概型：$P(A) = \dfrac{m}{n}$

4 ）条件概率：$P(B|A) = \dfrac{k}{m} = \dfrac{P(AB)}{P(A)}$

独立

5 ）乘法定理：$P(AB) = P(A)P(B|A) = P(A)P(B)$

$= 1 - P(\overline{A} \cup \overline{B})$

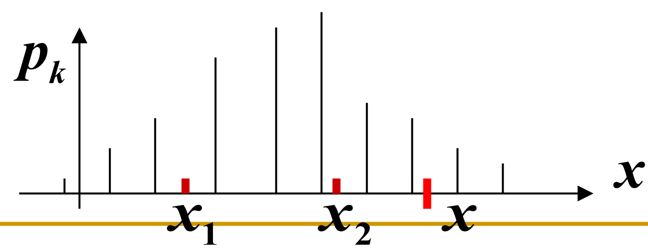
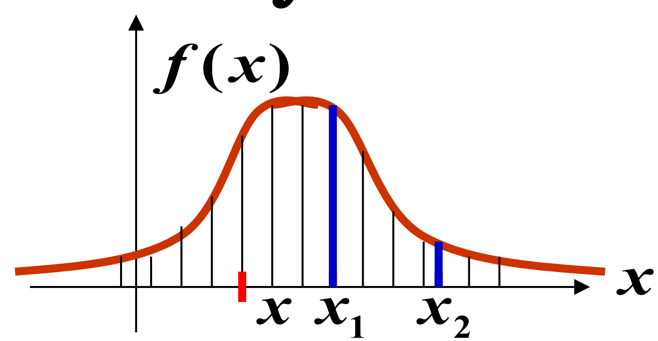
$A = AB_1 \cup AB_2$

6 ）全概率公式：$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2)$

7 ）贝叶斯公式：$P(B_1|A) = \dfrac{P(B_1)P(A|B_1)}{P(A)}$

$A \leftarrow B_1$，$A \leftarrow B_2$ 互斥

第二章 随机变量概率分布

	离散型随机变量	连续型随机变量
分布函数 $F(x)=P(X\le x)$ 概率的累加，不直观	$F(x)=\sum\limits_{x_k\le x} p_k$ 右连续	$F(x)=\int_{-\infty}^{x} f(t)dt$ 连续
概率分布 概率 1 分布情况，直观	分布律： $\sum p_k = 1$ $\begin{array}{c\|cccc} X & x_1 & x_2 & \cdots & x_k \\ \hline p_k & p_1 & p_2 & \cdots & p_k \end{array}$ 	概率密度： $\int_{-\infty}^{+\infty} f(t)dt = 1$ 
概率计算	$P(x_1<X\le x_2)=\sum\limits_{x_1<x_k\le x_2} p_k$ $=F(x_2)-F(x_1)$	$P(x_1<X\le x_2)=\int_{x_1}^{x_2} f(t)dt$ $=F(x_2)-F(x_1)$

非连续型随机变量

第二章

随机变量重要分布

	离散型随机变量	连续型随机变量
重要分布	1) (0-1) 分布 $P(X=k)=p^k(1-p)^{1-k}$	1) $U(a,b)$ $f(x)=\begin{cases}1/(b-a), & a<x<b\\ 0, & \text{其它}\end{cases}$
	2) $B(n,p)$★ $P(X=k)=C_n^k p^k(1-p)^{n-k}$	2) $E(\theta)$ $f(x)=\begin{cases}1/\theta\, e^{-x/\theta}, & x>0\\ 0, & \text{其它}\end{cases}$
	3) $P(\lambda)$ $P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$	3) $N(\mu,\sigma^2)$ ★ $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
函数的分布 $Y=g(X)$	X 的分布律 → Y 的分布律	★★ $f_X(x)\to f_Y(y)=F_Y'(y)$

第二章

1. 随机变量的分布函数 $F(x)=P(X\le x)$

作用： $P(x_1<X\le x_2)=F(x_2)-F(x_1)$

性质 1 $F(x)$ 是一个不减函数

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

性质 3 $F(x)$ 是右连续的函数

2. 连续型随机变量的概率密度 $F(x)=\int_{-\infty}^{x} f(t)dt$

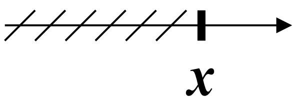
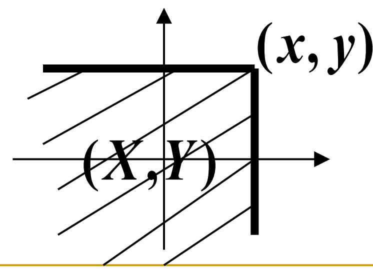
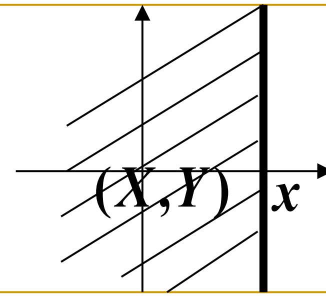
性质 1 、 2 $f(x)$ 0 $\int_{-\infty}^{\infty} f(x)dx=1$

性质 3 $P(x_1<X\le x_2)=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$

性质 4 $F'(x)=f(x)$

第三章 二维随机变量 (X,Y)

		(X,Y) 离散型	(X,Y) 连续型
(X,Y) 整体	联合分布函数 $F(x,y)$	联合分布律 $P(X=x_i,Y=y_j)=p_{ij}$	联合概率密度 $f(x,y)$
(X,Y) 个体	边缘分布函数 $F_X(x)=\lim\limits_{y\to\infty}F(x,y)$ $F_Y(y)=\lim\limits_{x\to\infty}F(x,y)$	边缘分布律 $P(X=x_i)=\sum\limits_{j=1}^{\infty}p_{ij}=p_{i\cdot}$ $P(Y=y_j)=\sum\limits_{i=1}^{\infty}p_{ij}=p_{\cdot j}$	边缘概率密度 $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy$ $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx$
X 与 Y 独立	对$\forall x,y$ $F(x,y)=F_X(x)F_Y(y)$	$P(X=x_i,Y=y_j)$ $=P(X=x_i)P(Y=y_j)$	$f(x,y)=f_X(x)f_Y(y)$
概率计算		$P\{(X,Y)\in G\}=\sum\limits_{(x_i,y_j)\in G}p_{ij}$	$P\{(X,Y)\in G\}$ $=\iint\limits_G f(x,y)dxdy$

	一维 X	二维 (X,Y)	边缘 X	关系
分布函数	$F(x)$ $=P(X\le x)$	$F(x,y)$ $=P(X\le x, Y\le y)$	$F_X(x)=P(X\le x)$ $=P(X\le x, Y<+\infty)$	$F_X(x)=\lim\limits_{y\to\infty}F(x,y)$
几何意义				第三章
离散型	$F(x)=\sum\limits_{x_k\le x} p_k$	$F(x,y)=\sum\limits_{\substack{x_i\le x\\ y_j\le y}} p_{ij}$	$F_X(x)=\sum\limits_{x_i\le x}\sum\limits_{j=1}^{\infty} P_{ij}$	
连续型	$F(x)$ $=\int_{-\infty}^{x} f(t)dt$	$F(x,y)$ ★ $=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv$	$F_X(x)$ $f_X(x)$ $=\int_{-\infty}^{x}\underline{\int_{-\infty}^{+\infty} f(x,y)dy}dx$	$f_X(x)$ ★ $=\int_{-\infty}^{+\infty} f(x,y)dy$
分布律	$P\{X=x_k\}=p_k$	$P\{X=x_i,Y=y_j\}$ $=p_{ij}$	$P\{X=x_i\}=\sum\limits_{j=1}^{\infty} p_{ij}$	$P\{X=x_i\}=\sum\limits_{j=1}^{\infty} p_{ij}$
概率	$P(x_1<X\le x_2)=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$		★ $P\{(X,Y)\in G\}=\iint\limits_G f(x,y)dxdx=\sum\limits_{(x_i,y_j)\in G} p_{ij}$	

第四节 两个随机变量的函数的分布

$Z=g(X,Y)$ $f(X,Y) \Rightarrow f_Z(z)=?$ $f_Z(z)=F'_Z(z)$

1) $Z=X+Y$ $f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$

独立

2) $Z=\max\{X,Y\}$ X,Y独立 $F_Z(z)=F_X(z)F_Y(z)$

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

$X_1,X_2,\cdots,X_n$独立，$F_{X_i}(x)=F(x)$

$M=\max(X_1,X_2,\cdots X_n)$ 与 $N=\min(X_1,X_2\cdots X_n)$

$F_{\max}(z)=F_{X_1}(z)\cdot F_{X_2}(z)\cdots F_{X_n}(z)=[F(z)]^n$

$F_{\min}(z)=1-[1-F_{X_1}(z)]\cdot[1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]=1-[1-F(z)]^n$

计算难点

1) $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv$

2) $P\{(X,Y)\in G\}=\iint\limits_{G} f(x,y)dxdx$

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

独立

4) $f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$

5) $Z=g(X,Y)$ $\quad f(X,Y)\longrightarrow f_Z(z)=?\quad f_Z(z)=F_Z'(z)$

$$F_Z(z)=P(Z\le z)=P\{g(X,Y)\le z\}=\iint\limits_{g(x,y)\le z} f(x,y)dxdy$$

$$=\iint\limits_{D(z)} f(x,y)dxdy$$

D是积分区域$g(x,y)\le z$与$f(x,y)$取值非零区域的交集

第四章 随机变量的数学期望与方差

	离散型随机变量	连续型随机变量
X	$E(X)=\sum_{k=1}^{\infty}x_k p_k$	$E(X)=\int_{-\infty}^{+\infty}xf(x)dx$
$Y=g(X)$ g 连续	$E(Y)=E[g(X)]$ $=\sum_{k=1}^{\infty}g(x_k)p_k$	$E(Y)=E[g(X)]$ $=\int_{-\infty}^{+\infty}g(x)f(x)dx$
$Z=g(X,Y)$ g 连续	$E(Z)=E[g(X,Y)]$ $=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}g(x_i,y_j)p_{ij}$	$E(Z)=E[g(X,Y)]$ $=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$
$D(X)$ $=E[X-E(X)]^2$	$D(X)=\sum_{k=1}^{\infty}(x_k-E(X))^2 p_k$	$D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2 f(x)dx$

第四章 随机变量的数字特征

$E(X)$ 性质	$E(c)=c \quad E(cX)=cE(X) \quad E(X+Y)=E(X)+E(Y)$ X,Y 独立 $\quad E(XY)=E(X)E(Y)$
$D(X)$ 性质	$D(X)=E(X^2)-[E(X)]^2 \quad D(c)=0 \quad D(cX)=c^2D(X)$ X,Y 独立 $D(X+Y)=D(X)+D(Y) \quad D(X)=0 \Leftrightarrow P(X=c)=1$
协方差	$Cov(X,Y)=E\{[X-E(X)][(Y-E(Y)]\}$ $=E(XY)-E(X)E(Y)$ $D(X+Y)=D(X)+D(Y)+2\operatorname{cov}(X,Y) \overset{\text{独立}}{=} D(X)+D(Y)$
相关系数	$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}$ (1). $\lvert\rho_{XY}\rvert\le 1$ (2). $\lvert\rho_{XY}\rvert=1 \Leftrightarrow$ 存在常数 a,b 使得：$P(Y=aX+b)=1$

第四章

几种常见分布的数学期望和方差

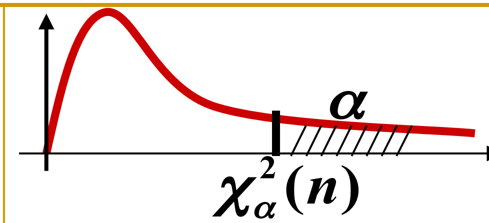
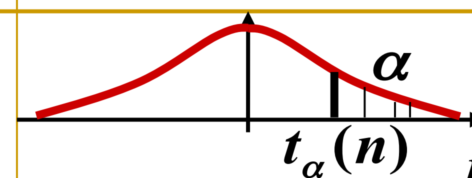
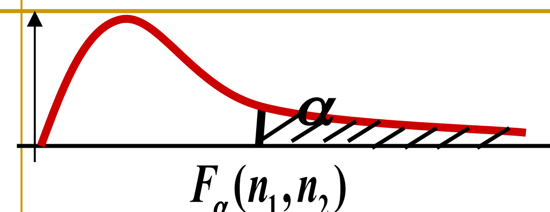
	概率分布		$E(X)$	$D(X)$
离散型	(0-1) 分布	$X \sim B(1,p)$	p	pq
	二项分布	$X \sim B(n,p)$	np	npq
	泊松分布	$X \sim P(\lambda)$	λ	λ
连续型	均匀分布	$X \sim U(a,b)$	$(a+b)/2$	$(b-a)^2/12$
	指数分布	$X \sim Exp(\theta)$	θ	θ^2
	正态分布	$X \sim N(\mu,\sigma^2)$	μ	σ^2

第五章 大数定律及中心极限定理

定理 1	$X_1, X_2, \cdots, X_n, \cdots$ 相互独立 $E(X_k)=\mu \quad D(X_k)=\sigma^2$	$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu$
定理 2 (贝努利)	$X_1, X_2, \cdots, X_n, \cdots$ 相互独立 $\sim (0-1)$分布(参数p)	$\frac{n_A}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} p$
定理 3 (辛钦)	$X_1, X_2, \cdots, X_n, \cdots$ 相互独立 $E(X_k)=\mu$ 同分布	$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu$
定理 1 (林德)	$X_1, X_2, \cdots, X_n, \cdots$ 相互独立 同分布 $E(X_k)=\mu \ D(X_k)=\sigma^2$	$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \overset{近似}{\sim} N(0,1)$ ★
定理 2 (德莫弗)	$X_n \sim B(n,p)$	$\frac{X_n - np}{\sqrt{np(1-p)}} \overset{近似}{\sim} N(0,1)$ ★

第六章

常用统计量及抽样分布

χ^2分布	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立 ★$\chi^2=\sum_{i=1}^{n}X_i^2 \sim \chi^2(n)$ $E(\chi^2)=n \quad D(\chi^2)=2n$	 $n>45$ $\chi^2_\alpha(n)\approx 1/2(z_\alpha+\sqrt{2n-1})^2$
t分布	$X\sim N(0,1),\ Y\sim\chi^2(n)$，独立 ★$t=\dfrac{X}{\sqrt{Y/n}}\sim t(n)$	 $n>45$ $t_\alpha(n)=-t_{1-\alpha}(n),\ t_\alpha(n)\approx z_\alpha$
F分布	$U\sim\chi^2(n_1), V\sim\chi^2(n_2)$，独立 ★$F=\dfrac{U/n_1}{V/n_2}\sim F(n_1,n_2)$ $1/F\sim F(n_2,n_1)$	 $F_{1-\alpha}(n_1,n_2)=1/F_\alpha(n_2,n_1)$
$X\sim N(\mu,\sigma^2)$ $X_1,X_2,\cdots,X_n$	★*Th*1 $\bar X\sim N(\mu,\sigma^2/n)$, $(n-1)S^2/\sigma^2\sim\chi^2(n-1)$ 独立	*Th*2 ★ $\dfrac{\bar X-\mu}{S/\sqrt n}\sim t(n-1)$

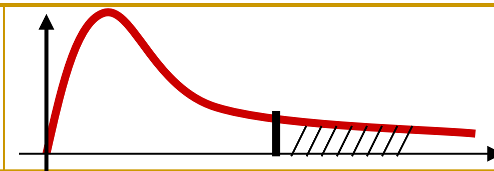
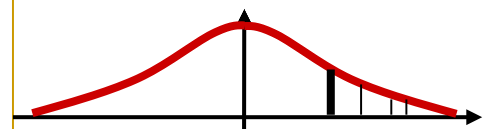
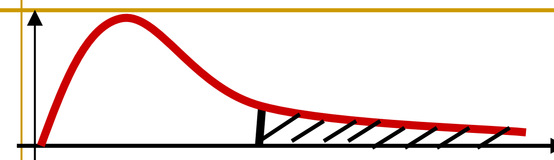
$\bar X, S^2 \qquad \bar X=\frac{1}{n}\sum_{i=1}^{n}X_i \qquad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar X)^2$

第六章

常用统计量及抽样分布

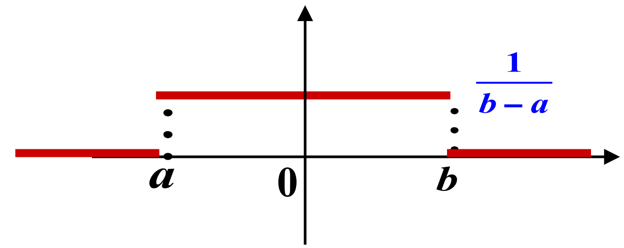
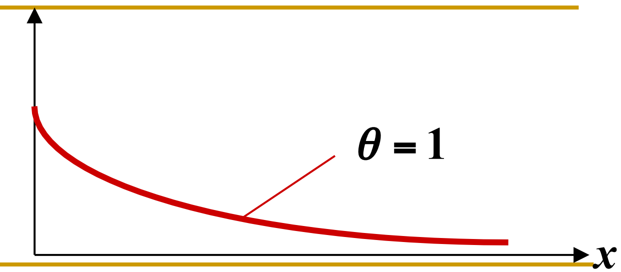
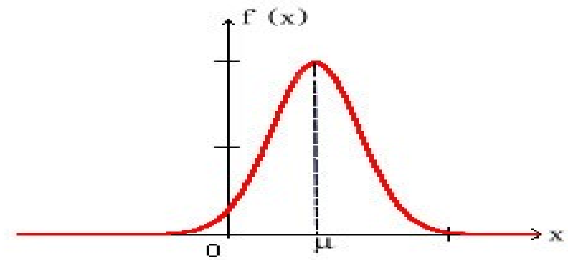
$X \sim N(\mu,\sigma^2)$

$X_1, X_2, \cdots, X_n$

χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$	
t 统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$	
F 统计量	$F = \dfrac{U/n_1}{V/n_2}$	$\sim F(n_1, n_2)$	
样本均值	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \dfrac{\sigma^2}{n})$	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\sim t(n-1)$	

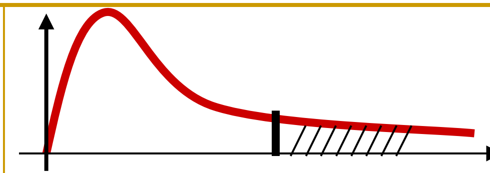
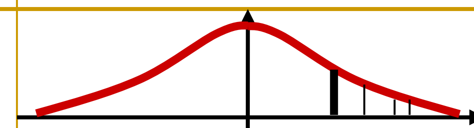
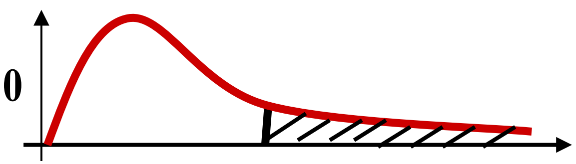
第六章

连续型随机变量及其分布

$X \sim U(a,b)$	$f(x)=\begin{cases}\dfrac{1}{b-a} & a<x<b \\ 0 & \text{其它}\end{cases}$	
$X \sim E(\theta)$	$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0 \\ 0 & \text{其它}\end{cases}$	
$X \sim N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	
$X \sim \Gamma(\alpha,\beta)$	$f(x)=\begin{cases}\dfrac{\beta^{\alpha_1}}{\Gamma(\alpha_1)}x^{\alpha_1-1}e^{-\beta x} & x>0 \\ 0 & x\le 0\end{cases}$	

第六章

常用统计量及抽样分布

$\chi^2 \sim \chi^2(n)$	$f(x)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(n/2)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & x \quad 0\\ 0 & x<0\end{cases}$	
$t \sim t(n)$	$t(x)=\dfrac{\Gamma[(n+1)/2]}{\Gamma(n/2)\sqrt{n\pi}}(1+\dfrac{x^2}{n})^{-\frac{n+1}{2}}$	
$F \sim F(n_1,n_2)$	$\varphi(x)=\begin{cases}\dfrac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})}(\frac{n_1}{n_2})(\frac{n_1}{n_2}x)^{\frac{n_1}{2}-1}\left(1+\frac{n_1}{n_2}x\right)^{-\frac{n_1+n_2}{2}}, & x \quad 0\\ 0 & x<0\end{cases}$	

第七章	总体 $X\sim F(x,\theta)$, $X_1X_2,\cdots,X_n$ $x_1x_2,\cdots,x_n$ 对θ进行估计
★ 点估计	统计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)\to\theta$估计量 1) 矩估计法：求解：$\mu_i=A_i,\ i=1,2,\cdots,k$ ★2) 极大似然估计法：求解：$L(\hat{\theta})=\max\limits_{\theta\in H}L(\theta)$
估计量的优良性	1) 无偏性：$E(\hat{\theta})=\theta$ 2) 有效性：$D(\hat{\theta}_1)<D(\hat{\theta}_2)$
★区间估计	$P(\underline{\theta}<\theta<\overline{\theta})=1-\alpha$ $(\underline{\theta},\overline{\theta})$是置信度为$1-\alpha$的置信区间 $X\sim N(\mu,\sigma^2)$，对μ,σ^2进行区间估计 1)求μ的置信区间σ^2为已知 2)求μ的置信区间σ^2为未知 3)求σ^2的置信区间

第七章

$X \sim N(\mu,\sigma^2)$，对μ,σ^2进行区间估计置信度$1-\alpha$

	统计量	置信区间
1)求μ的置信区间，σ^2为已知	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	★ $(\bar{X} \pm \dfrac{\sigma}{\sqrt{n}} z_{\alpha/2})$
2)求μ的置信区间，σ^2为未知	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	★ $(\bar{X} \pm \dfrac{S}{\sqrt{n}} t_{\alpha/2}(n-1))$
3)求σ^2的置信区间	$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	★ $\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$
(0-1) 分布 p 的置信区间	$\dfrac{n\bar{X}-np}{\sqrt{np(1-p)}} \sim N(0,1)$	(p_1, p_2) $p_{1,2} = \dfrac{1}{2a}(-b \mp \sqrt{b^2-4ac})$

第八章

$X \sim N(\mu,\sigma^2)$，对 μ,σ^2 进行假设检验显著性水平 α，

	原假设 H_0	备择假设 H_1	★ 检验统计量	★ 拒绝域
1) μ的检验 σ^2为已知	$\mu = \mu_0$ $\mu \le \mu_0$ $\mu \quad \mu_0$	$\mu \ne \mu_0$ $\mu > \mu_0$ $\mu < \mu_0$	$U = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ $\sim N(0,1)$	$\lvert U \rvert > z_{\alpha/2}$ $U > z_\alpha$ $U < -z_\alpha$
2) μ的检验 σ^2为未知	$\mu = \mu_0$ $\mu \le \mu_0$ $\mu \quad \mu_0$	$\mu \ne \mu_0$ $\mu > \mu_0$ $\mu < \mu_0$	$t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$ $\sim t(n-1)$	$\lvert t \rvert > t_{\alpha/2}(n-1)$ $t > t_\alpha(n-1)$ $t < -t_\alpha(n-1)$
3) σ^2的检验	$\sigma^2 = \sigma_0^2$ $\sigma^2 \le \sigma_0^2$ $\sigma^2 \quad \sigma_0^2$	$\sigma^2 \ne \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ $\sim \chi^2(n-1)$	$\chi^2 > \chi^2_{\alpha/2}(n-1)$或 $\chi^2 < \chi^2_{1-\alpha/2}(n-1)$ $\chi^2 > \chi^2_\alpha(n-1)$ $\chi^2 < \chi^2_{1-\alpha}(n-1)$