概率论与数理统计

第一章　随机事件

1. 概率的性质:
   (1) $P(\varnothing)=0$
   (2) A_i 两两互斥, $P(\bigcup A_i)=\sum P(A_i)$
   (3) $P(\overline{A})=1-P(A)$
   (4) $P(B-A)=P(B)-P(A)$
   (5) (加法公式) $P(A\cup B)=P(A)+P(B)-P(AB)$
2. **加法公式**: $P(A\cup B)=P(A)+P(B)-P(AB)$
   (1) 互不相容: $P(A\cup B)=P(A)+P(B)$
   (2) 相互独立: $P(A\cup B)=1-P(\overline{A})P(\overline{B})$
3. **乘法公式**: $P(AB)=P(A)P(B|A)=P(B)P(A|B)$
   (1) 相互独立: $P(AB)=P(A)P(B)$
4. **条件概率公式**: $P(B|A)=P_B(A)=\dfrac{P(AB)}{P(A)}$
5. **全概率公式**: $P(A)=\sum_{i=1}^n P(B_i)P(A|B_i)$
6. **贝叶斯公式**: $P(B_i|A)=\dfrac{P(B_i)P(A|B_i)}{\sum_{j=1}^n P(B_j)P(A|B_j)}$

第二章　随机变量

1. 概率密度函数 $f(x)$: $P\{a<X\leqslant b\}=\int_a^b f(x)\mathrm{d}x$
2. 概率分布函数: $F(x)=P\{X\leqslant x\}$
   (1) (单调不减性) $P\{a<X\leqslant b\}=F(b)-F(a)$
   (2) (有界性) $F(-\infty)=0, F(+\infty)=1$
   (3) $F(x)=\sum_{x_k\leqslant x}p_k$
   (4) $\sum F(x)=1$
3. $F(x)=P\{X\leqslant x\}=\int_{-\infty}^x f(t)\mathrm{d}t$
   $f(x)=F'(x)$
4. 已知连续性随机变量 X 的概率密度 $f_X(x)$, 求随机变量 $Y=g(X)$ 的概率密度 $f_Y(y)$
   (1) **积分转化法** $F_Y(y)=P\{g(X)\leqslant y\}$, $f_Y(y)=F'_Y(y)$
   (2) **单调函数公式法** 反函数 $x=h(y)$
   $$f_Y(y)=\begin{cases} f(h(y))\cdot|h'(y)|, & a<y<b,\\ 0, & \text{其他}\end{cases}$$

离散型随机变量

1. **两点分布** $X\sim B(1,p)$
   $\boxed{P\{X=1\}=p, P\{X=0\}=1-p}$
   性质 1 $E(X)=p, D(X)=p(1-p)$
   $\hat{p}=\bar{X}=\sum X/n$ 是 p 的矩估计和极大似然估计
2. **二项分布** $X\sim B(n,p)$ 伯努利实验
   $\boxed{b(k;n,p)=P\{X=k\}=\binom{n}{k}p^k(1-p)^{(n-k)}}$
   性质 1 $E(X)=np, D(X)=np(1-p)$
   性质 2 X 可表示成 n 个 $X_i\sim B(1,p)$ 的独立随机变量
   性质 3 若 X_i 相互独立, 且 $X_i\sim B(n_i,p)$,
   则 $X=\sum X_i\sim B(n,p)$
   $\hat{p}=\bar{X}/n=\sum X/(nN)$ 是 p 的矩估计和极大似然估计
3. **泊松分布** $X\sim P(\lambda)$
   $\boxed{b(k;\lambda)=P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}}$
   性质 1 $E(X)=\lambda, D(X)=\lambda$
   性质 2 若 X_i 相互独立, 且 $X_i\sim P(\lambda_i)$, 则 $X=\sum X_i\sim P(\lambda)$
   性质 3 若 X_1,X_2 相互独立, 且 $X_i\sim P(\lambda_i)$, 则 $X_i|(X_1+X_2=n)\sim B(n,\frac{\lambda_i}{\lambda_1+\lambda_2})$
   某医院每天前来就诊的病人数, 某地区一段时间间隔内发生火灾的次数, 一段时间间隔内容器内部细菌数, 某地一年内发生暴雨的次数, 每条床单上的疵点数.
4. $P\{X\geqslant x\}=\sum_{r=x}^{\infty}\frac{\mathrm{e}^{-r}\lambda^r}{r!}$
5. 当 $n\to\infty, p\to 0$ 时, $b(k;n,p)\approx b(k;\lambda), \lambda=np$

连续型随机变量

1. **均匀分布** $X\sim U[a,b]$
   $$\boxed{f(x)=\begin{cases}1/(b-a), & a\leqslant x\leqslant b,\\ 0, & \text{其他}\end{cases}}$$
   性质 1 $E(X)=(a+b)/2, D(X)=(b-a)^2/12$
   性质 2 $(\forall c,d)a\leqslant c<d\leqslant d$, 有 $P\{c\leqslant X\leqslant d\}=\dfrac{d-c}{b-a}$
   性质 3 $F(X)\sim U(0,1)$
2. **指数分布** $X\sim EP(\lambda)\sim e(\lambda^{-1})$ 寿命分布
   $$\boxed{f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x\geqslant 0,\\ 0, & x<0\end{cases}}$$
   性质 1 $E(X)=\lambda^{-1}, D(X)=\lambda^{-2}$
   性质 2 (无记忆性) $P\{X>x+y|X>y\}=P\{X>x\}$
3. **正态分布** $X\sim N(\mu,\sigma^2)$ 高斯分布
   $$\boxed{f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-u)^2}{2\sigma^2}}}$$
   (1) 标准正态分布: $X\sim N(0,1)$
   (2) 概率密度: $\varphi(x)\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$
   (3) 分布函数: $\Phi(x)=\int_{-\infty}^x\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u$
   (4) $\Phi(-x)=1-\Phi(x)$
   (5) 标准正态分布上的 α 分位点, $\Phi(Z_\alpha)=1-\alpha$,
   $P\{X>Z_\alpha\}=\int_{Z_\alpha}^{+\infty}\varphi(x)\mathrm{d}x=\alpha$
   性质 1 $E(X)=\mu, D(X)=\sigma^2$
   性质 2 $\sum_{i=1}^n c_iX_i+d\sim N\left(\sum_{i=1}^n c_i\mu_i+d, \sum_{i=1}^n c_i^2\sigma_i^2\right)$
   性质 3 若 $X\sim N(\mu,\sigma^2)$, 则 $Y=\frac{X-\mu}{\sigma}\sim N(0,1)$;
   若 $X\sim N(0,1)$, 则 $Y=\sigma X+\mu\sim N(\mu,\sigma^2)$
   性质 4 $\bar{X}=\frac{\sum_{i=1}^n X_i}{n}\sim N(\mu,\frac{\sigma^2}{n})$
   性质 5 $P\{a<X<b\}=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})$
   $P\{X<b\}=\Phi(\frac{b-\mu}{\sigma})$
   $P\{X>a\}=1-\Phi(\frac{a-\mu}{\sigma})$
   某地区成年男性的身高, 某零件长度的测量误差, 半导体器件中的热噪声电流.

第三章　随机向量

1. $F(x,y)=P\{X\leqslant x, Y\leqslant y\}$,
   $P\{X=x_i, Y=y_j\}=p_{ij}$
   (1) 离散型: $F(x,y)=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y}p_{ij}$
   (2) 连续型: $F(x,y)=\int_{-\infty}^y\int_{-\infty}^x f(u,v)\mathrm{d}u\mathrm{d}v$
   i. $f(x,y)\geqslant 0$
   ii. $F(+\infty,+\infty)=1$
   iii. $\dfrac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y)$
   iv. $P\{(X,Y)\in D\}=\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$
2. 二维均匀分布 $f(x,y)=\begin{cases}1/d, & (x,y)\in D,\\ 0, & \text{其他}\end{cases}$
3. 二维正态分布 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$
   $$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$
4. **边缘概率分布**:
   $F_X(x)=F(x,+\infty), F_Y(x)=F(+\infty,y)$
   (1) 二维离散型随机向量: $p_{i\cdot}=\sum_j p_{ij}$, $p_{\cdot j}=\sum_i p_{ij}$
   (2) 二维连续型随机变量: $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$
   $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$
5. **条件概率密度**: $f_{X|Y}(u|y)=\dfrac{f(x,y)}{f_Y(y)}$
6. 独立性的判断: $\forall x,y$, $F(x,y)=F_X(x)\cdot F_Y(y)$,
   对于连续型变量 $f(x,y)=f_X(x)\cdot f_Y(y)$
7. $Z=X+Y$ 的密度: $f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x$,
   $f_Z(z)=\int_{-\infty}^{+\infty}f(z-y,y)\mathrm{d}y$.

   (1) 相互独立: $f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x$,
   $f_Z(z)=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\mathrm{d}y$.
8. $Z=\max\{X,Y\}$ 和 $Z=\min\{X,Y\}$ 的分布:
   $F_{\max}(z)=F_X(z)\cdot F_Y(z)$
   $F_{\min}(z)=1-(1-F_X(z))\cdot(1-F_Y(z))$

第四章　数字特征

1. **期望** $E(X)$
   (1) 离散型: $E(X)=\sum_{i=1}^n x_ip_i$
   (2) 连续型: $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$
   (3) $E(g(X))=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x$
   性质 1 $E(c)=c$
   性质 2 $E(kX)=kE(X)$
   性质 3 $E(X+Y)=E(X)+E(Y)$
   性质 4 若 X,Y 相互独立, 则 $E(XY)=E(X)E(Y)$
2. **方差** $D(X)=\mathrm{Var}(X)=E\{[X-E(X)]^2\}$
   (1) 离散型: $D(X)=\sum_{i=1}^n[x_i-E(X)]^2p_i$
   (2) 连续型: $D(X)=\int_{-\infty}^{+\infty}[x_i-E(X)]^2f(x)\mathrm{d}x$
   (3) (计算公式) $D(X)=E(X^2)-E^2(X)$
   性质 1 $D(c)=0, D(X+c)=D(X)$
   性质 2 $D(kX)=k^2D(X), D(-X)=D(X)$
   性质 3 $D(X+Y)=D(X)+D(Y)+2\operatorname{Cov}(X,Y)$.
   若 X,Y 相互独立, 则 $D(X+Y)=D(X)+D(Y)$
   性质 4 若 X,Y 相互独立, 则 $D(aX\pm bY)=a^2D(X)+b^2D(Y)$
3. **协方差** $\operatorname{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$
   性质 1 $\operatorname{Cov}(X,Y)=\operatorname{Cov}(Y,X)$
   性质 2 $\operatorname{Cov}(aX+b,cY+d)=ac\operatorname{Cov}(X,Y)$
   性质 3 $\operatorname{Cov}(X_1+X_2,Y)\operatorname{Cov}(X_1,Y)+\operatorname{Cov}(X_2,Y)$
   性质 4 (计算公式) $\operatorname{Cov}(X,Y)=E(XY)-E(X)E(Y)$
   性质 5 $\operatorname{Cov}^2(X,Y)\leqslant D(X)D(Y)$, 等号成立 iff $(\exists a,b)Y=aX+b$
4. **相关系数** $\rho_{XY}=\dfrac{\operatorname{Cov}(X,Y)}{\sqrt{D(X)D(Y)}}$
   (1) $|\rho_{XY}|\leqslant 1$
   (2) 互不相关: $\rho_{XY}=0$
5. $E\{[X-E(X)]^k\}$ 称为 X 的 k 阶**中心矩**
6. 记 $c_{ij}=\operatorname{Cov}(X_i,Y_j)$, 称 $\boldsymbol{C}=(c_{ij})_{n\times n}$ 为 $(X_1,X_2,\cdots,X_n)$ 的**协方差矩阵**

第五章　极限定理

1. **切比雪夫不等式**: $\forall\varepsilon>0$,
   $P\{|X-\mu|\geqslant\varepsilon\}\leqslant\sigma^2/\varepsilon^2$,
   $P\{|X-\mu|<\varepsilon\}\geqslant 1-\sigma^2/\varepsilon^2$
2. **大数定律**: X_i 相互独立, $Y_n=\frac{1}{n}\sum_{i=1}^n X_i$, $\forall\varepsilon>0$,
   $\lim_{n\to\infty}P\{|Y_n-\mu|<\varepsilon\}=1$.
   $\lim_{n\to\infty}P\{|n_A/n-p|<\varepsilon\}=1$.
3. 独立同分布的**中心极限定理**:
   X_i 相互独立且服从同一分布, 则 $Y_n=\dfrac{\sum_{i=1}^n X_i-n\mu}{\sqrt{n}\sigma}$
   的分布函数 $F_n(x)=P\{Y_n\leqslant x\}$ 收敛于 $\Phi(x)$,
   即 $\lim_{x\to\infty}F_n(x)=\Phi(x)$.
4. (棣莫夫-拉普拉斯定理) X_i 相互独立且服从 $B(1,p)$, 则 $\forall x$,
   $$\lim_{x\to\infty}P\left\{\frac{\sum_{i=1}^n X_i-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\Phi(x).$$

第六章　样本与统计量

1. **样本均值** $\bar{X}=\frac{1}{n}\sum_{i=1}^n X_i$ 是 μ 的估计
2. **样本方差** $S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$ 是 σ^2 的估计
3. 样本标准差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2}$
4. **样本原点矩** $A_k=\frac{1}{n}\sum_{i=1}^n X_i^k$
5. **样本中心矩** $M_k=\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^k$
6. 样本均值的**大样本分布**: $\bar{X}\sim N(\mu,\dfrac{\sigma^2}{n})$.

$F(x) \approx \Phi\Big(\dfrac{x-\mu}{\sigma/\sqrt{n}}\Big)$,

$P\{|\bar{X}-\mu| \leqslant c\} \approx 2\Phi\Big(\dfrac{c}{\sigma/\sqrt{n}}\Big)-1$

正态总体的抽样分布

1. $\Gamma\Big(\dfrac{n}{2}\Big) = \begin{cases} (\frac{n}{2}-1)(\frac{n}{2}-2)\cdots 3\cdot 2\cdot 1, & n\text{为偶数}, \\ (\frac{n}{2}-1)(\frac{n}{2}-2)\cdots \frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi}, & n\text{为奇数}. \end{cases}$

2. χ^2 **分布** $X \sim \chi^2_n$

   X_i 独立同分布于 $N(0,1)$, $X = \sum_{i=1}^n X_i^2$

   $$\boxed{f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}, & x>0, \\ 0, & x \leqslant 0. \end{cases}}$$

   性质1 $E(X^k) = 2^k \frac{\Gamma(\frac{n}{2}+k)}{\Gamma(\frac{n}{2})}$,
   $E(X) = n, D(X) = 2n$

   性质2 若 $X_i \sim \chi^2_{n_i}$ 独立同分布, 则 $\sum_{i=1}^m X_i \sim \chi^2_n$,
   其中 $n = \sum_{i=1}^m n_i$

3. $\chi^2_n(\alpha)$ 为 χ^2_n 分布上的 α 分位点:

   $P\{\chi^2_n > \chi^2_n(\alpha)\} = \int_{\chi^2_n(\alpha)}^{+\infty} f(x)\mathrm{d}x = \alpha$

4. t **分布** $T \sim t_n$ 学生分布

   $X \sim N(0,1), Y \sim \chi^2_n, T = \dfrac{X}{\sqrt{Y/n}}$

   $$\boxed{f(x) = \dfrac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\Big(1+\dfrac{x^2}{n}\Big)^{-\frac{n+1}{2}}}$$

   性质1 $E(F) = 0, D(F) = n/(n-2), n = 3, 4, \cdots$

   性质2 $Y = \dfrac{F_2 - F_1}{2\sqrt{F_1F_2/n}} \sim t_n$

   性质3 $\dfrac{\overline{F}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

5. $t_n(\alpha)$ 为 t_n 分布上的 α 分位点:

   $P\{T > t_n(\alpha)\} = \int_{t_n(\alpha)}^{+\infty} f(x)\mathrm{d}x = \alpha$

6. (基本定理) 样本 X_i 来自正态总体 $N(\mu,\sigma^2)$
   (1) $\bar{X} \sim N(\mu, \sigma^2/n)$
   (2) $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$
   (3) $\bar{X}$ 与 S^2 相互独立
   (4) $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

第七章　参数估计

1. 矩估计: $\alpha_m = E(X^m), A_m = \sum X_i^m, \alpha_m(\theta_i) = A_m$
   (1) 正态分布: $\hat{\mu} = \bar{X}, \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$
   (2) 均匀分布: $\hat{a} = \bar{X} - \sqrt{3}\hat{\sigma}, \hat{b} = \bar{X} + \sqrt{3}\hat{\sigma}$
2. 极大似然估计: $L(x_i;\theta_j) = \prod f(x_i, \theta_j)$,
   $\ln L(\theta) = \sum \ln f(x_i;\theta_j)$, $\frac{\partial \ln L(\theta)}{\partial \theta_i} = 0$
   (1) 正态分布: $\mu^* = \bar{X}, \sigma^{2*} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$
   (2) 泊松分布: $\lambda^* = \bar{X}$
   (3) 均匀分布: $a^* = \min\{X_i\}, b^* = \max\{X_i\}$
3. 无偏估计: $E[\hat{\theta}(X_i)] = \theta$
   (1) $E(\bar{X}) = \mu$
   (2) $E(S^2) = \sigma^2$
   (3) S 不是 σ 的无偏估计
4. 均方误差: $\mathrm{MSE}(\hat{\theta}) = E[(\hat{\theta}-\theta)^2] = D(\hat{\theta}) + [E(\hat{\theta})-\theta]^2$
   (1) 无偏估计: $\mathrm{MSE}(\hat{\theta}) = D(\hat{\theta})$
   (2) 两个估计中哪个方差小, 哪个优
5. $P\{\theta \in [\hat{\theta}_1, \hat{\theta}_2]\} \geqslant 1-\alpha$
   置信区间 $[\hat{\theta}_1, \hat{\theta}_2]$, 置信系数 $1-\alpha$
6. $X_i \sim N(\mu, \sigma^2)$, 置信系数为 $1-\alpha$
   (1) μ 的置信区间: $[\bar{X} - \frac{\sigma}{\sqrt{n}}Z_{\frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}}Z_{\frac{\alpha}{2}}]$
   σ^2 未知时: $[\bar{X} - \frac{S}{\sqrt{n}}t_{n-1}(\frac{\alpha}{2}), \bar{X} + \frac{S}{\sqrt{n}}t_{n-1}(\frac{\alpha}{2})]$
   (2) σ^2 的置信区间: $[\frac{(n-1)S^2}{\chi^2_{n-1}(\frac{\alpha}{2})}, \frac{(n-1)S^2}{\chi^2_{n-1}(1-\frac{\alpha}{2})}]$

7. $X_i \sim N(\mu_1, \sigma_1^2), Y_i \sim N(\mu_2, \sigma_2^2)$ 互相独立
   (1) $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \sigma_1^2/m + \sigma_2^2/n)$,
   或 $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/m+\sigma_2^2/n}} \sim N(0,1)$
   (2) $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S\sqrt{1/m+1/n}} \sim t_{m+n-2}$,
   其中 $S^2 = \frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}$
   (3) $\mu_1 - \mu_2$ 的置信区间: $[\bar{X} - \bar{Y} \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{\sqrt{m}} + \frac{\sigma_2^2}{\sqrt{n}}}]$
   $\sigma_1^2 = \sigma_2^2 = \sigma^2$: $[\bar{X} - \bar{Y} \pm t_{m+n-2}(\frac{\alpha}{2})S\sqrt{\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}}}]$
8. 大样本参数 μ 的置信区间: $[\bar{X} \pm \frac{\sigma}{\sqrt{n}}Z_{\frac{\alpha}{2}}]$
9. 二项分布参数 p 的置信区间: $[\hat{p} \pm Z_{\frac{\alpha}{2}}\sqrt{\hat{p}(1-\hat{p})/n}]$,
   其中 $\hat{p} = Y_n/n = \sum_{i=1}^n X_i/n$
10. 泊松分布参数 λ 的置信区间: $[\bar{X} \pm Z_{\frac{\alpha}{2}}\sqrt{\bar{X}/n}]$

第八章　假设检验

1. 检验 H_0 有可能犯的两类错误
   (1) 第一类错误“弃真”错误: H_0 正确但是被拒绝了
   (2) 第二类错误“采伪”错误: H_0 错误但是被接受了
2. 检验的显著性水平: $P\{\text{犯第一类错误}\} \leqslant \alpha \in (0,1)$

1.　正态总体均值的检验

1. 单个正态总体 $N(\mu, \sigma^2)$ 均值 μ 的检验
   (1) $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu \neq \mu_0$
      i. 拒绝域: $\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geqslant Z_{\frac{\alpha}{2}}$
      ii. 拒绝域: $\frac{|\bar{X}-\mu_0|}{S/\sqrt{n}} \geqslant t_{n-1}(\frac{\alpha}{2})$
   (2) $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu > \mu_0$
      i. 拒绝域: $\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \geqslant t_{n-1}(\alpha)$
2. 两个正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 均值的比较
   (1) $H_0: \mu_1 = \mu_2 \leftrightarrow H_1: \mu_1 \neq \mu_2$
      i. 拒绝域: $|\bar{X} - \bar{Y}| \geqslant Z_{\frac{\alpha}{2}}\sqrt{\sigma_1^2/m + \sigma_2^2/n}$
      ii. 拒绝域: $|\bar{X} - \bar{Y}| \geqslant t_{m+n-2}(\frac{\alpha}{2})\sqrt{\frac{m+n}{mn}}S$
   (2) $H_0': \mu_1 - \mu_2 \geqslant 0 \leftrightarrow H_1': \mu_1 - \mu_2 < 0$
      i. 拒绝域: $\bar{X} - \bar{Y} \leqslant -t_{m+n-2}(\alpha)\sqrt{\frac{m+n}{mn}}S$
   (3) $H_0'': \mu_1 - \mu_2 \leqslant 0 \leftrightarrow H_1'': \mu_1 - \mu_2 > 0$
      i. 拒绝域: $\bar{X} - \bar{Y} \geqslant t_{m+n-2}(\alpha)\sqrt{\frac{m+n}{mn}}S$
3. 成对数据的 t 检验
   (1) $H_0: \mu = 0 \leftrightarrow H_1: \mu \neq 0$, $d_i = X_i - Y_i$
      i. 拒绝域: $|\bar{d}| \geqslant t_{n-1}(\frac{\alpha}{2})\frac{S_d}{\sqrt{n}}$

2.　正态总体方差的检验

1. 单个正态总体方差的 χ^2 检验
   (1) $H_0: \sigma^2 = \sigma_0^2 \leftrightarrow H_1: \sigma^2 \neq \sigma_0^2$
      i. 拒绝域: $(n-1)S^2/\sigma_0^2 \leqslant \chi^2_{n-1}(1-\frac{\alpha}{2})$
      ii. 拒绝域: $(n-1)S^2/\sigma_0^2 \geqslant \chi^2_{n-1}(\frac{\alpha}{2})$
   (2) $H_0: \sigma^2 \leqslant \sigma_0^2 \leftrightarrow H_1: \sigma^2 > \sigma_0^2$
      i. 拒绝域: $(n-1)S^2/\sigma_0^2 \geqslant \chi^2_{n-1}(\alpha)$

3.　拟合优度检验

1. H_0 :样本 X_i 的总体分布为 $F(x)$
   (1) 把 $(-\infty, +\infty)$ 分割为 k 个区间:
   $I_1 = (a_0, a_1], \cdots, I_k = (a_{k-1}, a_k)$
   (2) 计算每个区间的实际频数 f_i
   (3) 计算理论频数 $p_i(\theta) = F(a_i, \theta) - F(a_{i-1}, \theta)$
   (4) 计算偏差平方和 $\chi^2 = \sum_{i=1}^k \frac{[f_i - np_i(\theta^*)]^2}{np_i(\theta^*)}$
   (5) 假设 H_0 的水平 α 的检验拒绝域为 $\chi^2 \geqslant \chi^2_{k-r-1}(\alpha)$

4.　独立性检验

1. $H_0: p_{ij} = p_{i\cdot}p_{\cdot j}$
   (1) 拒绝域: $\chi^2 \geqslant \chi^2_{(a-1)(b-1)}(\alpha)$,
   其中 $\chi^2 = \sum_{i=1}^a \sum_{j=1}^b \frac{(nn_{ij} - n_{i\cdot}n_{\cdot j})^2}{n_{ij}n_{i\cdot}n_{\cdot j}}$

第九章　回归分析

1. 一元线性回归模型 $y_i = \beta_0 + \beta_1 x_i + e_i$
   性质1 $E(\hat{\beta}_0) = \beta_0, E(\hat{\beta}_1) = \beta_1$ 为无偏估计

   性质2 假设 $e_i \sim N(0, \sigma^2)$, $S_{xx} = \sum(x_i - \bar{x})^2$,则
   $\hat{\beta}_0 \sim N(\beta_0, (1/n + \bar{x}^2/S_{xx})\sigma^2)$,
   $\hat{\beta}_1 \sim N(\beta_1, \sigma^2/S_{xx})$.
2. 高斯-马尔可夫假设: $E(e_i) = 0$,$D(e_i) = \sigma^2, \mathrm{Cov}(e_i, e_j) = 0$
3. $L_{xx} = \sum x_i^2 - n\bar{x}^2$,
   $L_{xy} = \sum x_i y_i - n\overline{xy}$,
   $L_{yy} = \sum y_i^2 - n\bar{y}^2$.
4. 最小二乘估计 $\begin{cases} \hat{a} = \bar{y} - \hat{b}\bar{x}, \\ \hat{b} = \frac{\sum x_i y_i - n\overline{xy}}{\sum x_i^2 - n\bar{x}^2} = \frac{L_{xy}}{L_{xx}}. \end{cases}$
5. 经验回归直线方程 $\hat{y} = \hat{b}x + \hat{a}$
6. 残差 $\hat{e}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$
7. $\hat{\sigma}^2 = \sum \hat{e}_i^2/(n-2)$
   (1) $\hat{\sigma}^2$ 是 σ^2 的无偏估计
   (2) $(n-2)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n-2}$,且 $\hat{\sigma}^2$ 与 $\hat{\beta}_0, \hat{\beta}_1$ 相互独立
8. 回归方程的显著性检验
   (1) $H_0: \beta_1 = 0 \leftrightarrow H_1: \beta_1 \neq 0$
      i. 拒绝域: $r = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}} \geqslant r_{\alpha=0.01}(n-2)$
      ii. 拒绝域: $|T| = \Big|\frac{\hat{\beta}_1^2}{\hat{\sigma}/\sqrt{S_{xx}}}\Big| \geqslant t_{n-2}(\frac{\alpha}{2})$
      iii. 拒绝域: $F = \frac{\hat{\beta}_1^2}{\hat{\sigma}^2/S_{xx}} \geqslant F_{1,n-2}$
9. 判定系数/确定系数 $R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$
   (1) 回归平方和 $SS_{\text{回}} = \sum(\hat{y}_i - \bar{y})^2$
   (2) 总平方和 $SS_{\text{总}} = \sum(y_i - \bar{y})^2$
   (3) 误差平方和 $SS_{\text{误}} = \sum \hat{e}_i^2 = SS_{\text{回}} - SS_{\text{总}}$
10. 回归参数的区间估计 $\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}\hat{\sigma}$, $\hat{\sigma}_{\hat{\beta}_1} = \frac{\hat{\sigma}}{\sqrt{S_{xx}}}$.
   β_i 的置信区间: $[\hat{\beta}_i \pm t_{n-2}(\frac{\alpha}{2})\hat{\sigma}_{\hat{\beta}_i}]$

常用结论

1. 人群中有相同生日: $P(A) = \frac{A_N^n}{N^n}$
2. 把 n 个物品分成 k 组, 每组恰有 n_i 个, 不同分组方法 $\frac{n!}{\prod_{i=1}^k n_i!}$ 种
3. 将 n 个球放入 M 个盒子, 有球的盒子数 $E(X) = M\left[1 - (1 - \frac{1}{M})^n\right]$
4. **基本积分表**
   (1) $\int k\mathrm{d}x = kx + C$
   (2) $\int x^\mu \mathrm{d}x = \frac{x^{\mu+1}}{\mu+1} + C$
   (3) $\int \frac{\mathrm{d}x}{x} = \ln|x| + C$
   (4) $\int \frac{\mathrm{d}x}{1+x^2} = \arctan x + C$
   (5) $\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \arcsin x + C$
   (6) $\int \cos x\mathrm{d}x = \sin x + C$
   (7) $\int \sin x\mathrm{d}x = -\cos x + C$
   (8) $\int \mathrm{e}^x\mathrm{d}x = \mathrm{e}^x + C$
   (9) $\int a^x\mathrm{d}x = \frac{a^x}{\ln a} + C$
5. 微分法则
   $\mathrm{d}(u \pm v) = \mathrm{d}u \pm \mathrm{d}v$
   $\mathrm{d}(Cu) = C\mathrm{d}u$
   $\mathrm{d}(uv) = v\mathrm{d}u + u\mathrm{d}v$
   $\mathrm{d}(\frac{u}{v}) = \frac{v\mathrm{d}u - u\mathrm{d}v}{v^2}$
6. 积分换元法
   $\int f[\varphi(x)]\varphi'(x)\mathrm{d}x = [\int f(u)\mathrm{d}u]_{u=\varphi(x)}$
   $\int f(x)\mathrm{d}x = [f[\psi(t)]\psi'(t)\mathrm{d}t]_{t=\psi^{-1}(x)}$
7. 二重积分的计算
   $\iint_D f(x,y)\mathrm{d}\sigma = \int_a^b \mathrm{d}x\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\mathrm{d}y$
   $\iint_D f(x,y)\mathrm{d}\sigma = \int_c^d \mathrm{d}y\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\mathrm{d}x$