做了全部的期末历年题和四套综合测试题，我把觉得自己做错了的和比较有意义的题目写了出来，前面是一些琐碎知识点，复习过程中我看的笔记是一位学长的，感觉非常详细，排版也很好看，点这里

# 第一章 基本概念

$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$

德摩根定律：

$$\overline{\bigcup_{n\geq 1} An}=\bigcap_{n\geq 1}\overline{An}$$

$$\overline{\bigcap_{n\geq 1} An}=\bigcup_{n\geq 1}\overline{An}$$

贝叶斯公式：

如果$S$中对事件$B$的假定构成了互斥、完备的一组分划：$A_1,A_2,...,A_n$，，且各事件概率均为正，那么可得：

$$P(A_i|B)=\frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n}P(B|A_j)P(A_j)}.$$

独立：如果A与B独立，则A非与B非独立，A与B非独立，A非与B独立

事件A与事件B不相容指的就是A与B互斥

# 第二章 随机变量及其概率分布

**1. 0-1分布：**
$X\sim 0-1(p)$或者$X\sim B(1,p)$
$E(X)=p$
$Var(X)=p(1-p)$

**2. 二项分布：**
$X\sim B(n,p)$

$E(X) = np$
$Var(X) = np(1-p)$

**3. 泊松分布:**
概率分布律（λ应为正数）:

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

$X \sim P(\lambda)$
$E(X) = \lambda$
$Var(X) = \lambda$

**4. 均匀分布:**
从a到b的均匀分布
$X \sim U(a,b)$
$E(X) = \frac{a+b}{2}$
$Var(X) = \frac{(b-a)^2}{12}$

**5. 正态分布:**
密度函数:

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

$X \sim N(a,b)$
$E(X) = \mu$
$Var(X) = \sigma^2$

**关于正态分布:**
如果A和B是相互独立的正态分布，那么A和B的线性组合也是正态分布
如果A和B是联合正态分布（或者说服从二元正态分布），那么不管A和B是否相互独立，A和B的线性组合都服从正态分布，但是要注意，如果不相互独立，协方差不为0
也就是下面这个公式:

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

如果是X-Y的话协方差前面就应该是负号

**6. 指数分布：**
密度函数：

$$f(x|\lambda) = \lambda e^{-\lambda x}$$

当x小于0的时候是0，要不然指数爆炸了
分布函数：

$$F(x|\lambda) = 1 - e^{-\lambda x}$$

$E(X) = \frac{1}{\lambda}$
$Var(X) = \frac{1}{\lambda^2}$
指数分布的特点是无记忆性：

$$P(X > t_0 + t \mid X > t_0) = \frac{P(X > t_0 + t)}{P(X > t_0)} = \frac{1 - F(t_0 + t)}{1 - F(t_0)} = e^{-\lambda t} = P(X > t)$$

# 随机变量的函数分布

函数)，记 $y = g(x)$ 的反函数为 $x = h(y)$，则 $Y$ 的密度函数为

$$f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)|, & y \in D, \\ 0, & y \notin D, \end{cases}$$

其中 $D$ 为函数 $y = g(x)$ 的值域.

上面这个方法成立的前提是，$y = g(x)$是一个处处可导的严格单调递增函数或者严格单调递减函数，也就是要做到一一对应

下面这个图和上面是一样的，更清晰一点：

特别地，当 $y=g(x)$ 具有**严格单调性**（严格单增 / 严格单减）时，记 $X$ 的密度函数为 $f_X(x)$，$y=g(x)$ 的反函数为 $x=h(y)$，则 $Y$ 的密度函数为：

$$f_Y(y)=\begin{cases}f_X(h(y))\cdot|h'(y)| & ,y\in D\\ 0 & ,y\notin D\end{cases}$$

其中 $D$ 为函数 $y=g(x)$ 的值域。

# 第三章 多维随机变量及其分布

以上结果容易推广到 $n$ 个变量的情形. 特别地, 设 $X_1,X_2,\cdots,X_n$ 为 $n$ 个相互独立的随机变量, 相应的分布函数分别为 $F_1(x),F_2(x),\cdots,F_n(x)$, 记 $M=\max\{X_1,X_2,\cdots,X_n\}$, $N=\min\{X_1,X_2,\cdots,X_n\}$, 则

$$F_M(t)=\prod_{i=1}^{n}F_i(t), \tag{3.5.15}$$

$$F_N(t)=1-\prod_{i=1}^{n}[1-F_i(t)]. \tag{3.5.16}$$

注意前提条件是相互独立

$$\begin{aligned}f_{Y|X}(y|x)&=\frac{f(x,y)}{f_X(x)}\\&=\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\{\frac{-1}{2(1-\rho^2)\sigma_2^2}[y-(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))]^2\end{aligned}$$

## 二元正态分布(二维正态变量)

对于二元正态分布的X和Y的线性组合，如果这些线性组合不相关，则可以直接推出独立

$$(X,Y)\ \ N(\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2,\rho)$$

## MAX和MIN问题

推广到 $n$ 个变量：设 $X_1,X_2,\ldots,X_n$ 为 $n$ 个相互独立的随机变量，分布函数分别为 $F_1(x),F_2(x),\ldots,F_n(x)$，记 $M=\max\{X_1,X_2,\ldots,X_n\},N=\min\{X_1,X_2,\ldots,X_n\}$，则

$$F_M(t)=\prod_{i=1}^{n}F_i(t)$$

$$F_N(t)=1-\prod_{i=1}^{n}[1-F_i(t)]$$

要注意，分布函数没有要求是一样的，这些相互独立的随机变量的分布函数可以不一样

# 第四章 随机变量的数字特征

前提是相互独立：

$W=XY$, 且 $X$ 与 $Y$ 相互独立, 故

$$E(W)=E(XY)=E(X)\cdot E(Y)$$

对于两两独立的随机变量：
方差可以直接相加

对于两个不独立的随机变量：
方差相加，还要加上两倍的协方差

Var(x)和D(x)都是方差

- $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
  - 可以看出，$\dfrac{S}{\sigma} \sim N(0,1)$

$E(S^2) = \delta^2$

协方差的性质:

**定理 4.3.2** 若随机变量 $X$ 和 $Y$ 的协方差存在, 则

(1) $\operatorname{Cov}(X, Y) = \operatorname{Cov}(Y, X)$;

(2) $\operatorname{Cov}(X, X) = \operatorname{Var}(X)$;

(3) $\operatorname{Cov}(aX, bY) = ab\operatorname{Cov}(X, Y)$, 其中 $a, b$ 为两个实数;

(4) 若 $\operatorname{Cov}(X_i, Y)(i = 1, 2)$ 存在, 则

$$\operatorname{Cov}(X_1 + X_2, Y) = \operatorname{Cov}(X_1, Y) + \operatorname{Cov}(X_2, Y);$$

(5) 若 $X$ 和 $Y$ 相互独立, 则 $\operatorname{Cov}(X, Y) = 0$, 但反之不然;

(6) 当 $\operatorname{Var}(X) \cdot \operatorname{Var}(Y) \neq 0$ 时, 有 $(\operatorname{Cov}(X, Y))^2 \leqslant \operatorname{Var}(X)\operatorname{Var}(Y)$, 其中等号成立当且仅当 $X$ 与 $Y$ 之间有严格的线性关系 (即存在常数 $c_1, c_2$ 使得 $P\{Y = c_1 + c_2 X\} = 1$ 成立).

通过前面的性质，可以证明性质6

相关系数表达式:

$$r = \frac{\operatorname{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

独立-->不相关

**独立一定不相关，不相关不一定独立**

**相关一定不独立，不独立不一定相关**

# 第五章 大数定律及中心极限定理

切比雪夫不等式不针对某一特定的分布

**切比雪夫（ChebyShev）不等式**：设随机变量 $X$ 的数学期望和方差存在，分别记为 $\mu, \sigma^2$，则对任意的 $\varepsilon > 0$，有：

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad (\text{or } P(|X-\mu| < \varepsilon) \geq 1-\frac{\sigma^2}{\varepsilon^2})$$

辛钦大数定律**只要求期望存在，不要求方差存在**

下面这两个中心极限定律解决的是求和的问题（一个是用来求独立同分布的随机变量序列的和，一个是用来求n重伯努利实验）

**林德伯格（Lindeberg）- 莱维（Lévy）中心极限定理**：设 $\{X_i, i \geq 1\}$ 为独立同分布的随机变量序列，且期望 $E(X_i)=\mu$ 和方差 $Var(X_i)=\sigma^2$ 均存在 $(\sigma > 0)$，则对任意的 $x \in \mathbf{R}$，有：

$$\lim_{n\to+\infty} P\left(\frac{\sum_{i=1}^{n} X_i - E(\sum_{i=1}^{n} X_i)}{\sqrt{Var(\sum_{i=1}^{n} X_i)}} \leq x\right) = \lim_{n\to+\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq x\right)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\mathrm{d}t = \Phi(x)$$

- 该定理表明：期望为 $\mu$，方差为 $\sigma^2$ 的独立同分布的速记变量的部分和 $\sum_{i=1}^{n} X_i$ 的标准化变量 $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$，当 $n$ 充分大时，近似地服从**标准正态分布** $N(0,1)$，即：

**棣莫弗（De Moivre）- 拉普拉斯（Laplace）中心极限定理**：设 $n_A$ 为 $n$ 重伯努利试验中事件 $A$ 发生的次数，$p$ 为事件 $A$ 在每次试验中发生的概率，即 $P(A)=p$，则对任意 $x\in\mathbf{R}$，有：

$$\lim_{n\to+\infty}P\left(\frac{n_A-np}{\sqrt{np(1-p)}}\leq x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\mathrm{d}t=\Phi(x)$$

- 该定理表明，当 $n$ 充分大时，二项分布 $B(n,p)$ 可用正态分布 $N(np,np(1-p))$ 来逼近。

# 第七章 参数估计

| | 待估参数 | 其他参数 | 枢轴量的分布 | 置信区间 | 单侧置信限 |
|---|---|---|---|---|---|
| 单个正态总体 | $\mu$ | $\sigma^2$ 已知 | $Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ | $\left(\overline{X}\pm\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$ | $\hat{\mu}_U=\overline{X}+\frac{\sigma}{\sqrt{n}}z_\alpha$ $\quad\hat{\mu}_L=\overline{X}-\frac{\sigma}{\sqrt{n}}z_\alpha$ |
| | $\mu$ | $\sigma^2$ 未知 | $t=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$ | $\left(\overline{X}\pm\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$ | $\hat{\mu}_U=\overline{X}+\frac{S}{\sqrt{n}}t_\alpha(n-1)$<br>$\hat{\mu}_L=\overline{X}-\frac{S}{\sqrt{n}}t_\alpha(n-1)$ |
| | $\sigma^2$ | $\mu$ 未知 | $\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ | $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$ | $\hat{\sigma}^2_U=\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$ $\quad\hat{\sigma}^2_L=\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}$ |
| 两个正态总体 | $\mu_1-\mu_2$ | $\sigma_1^2$ 和 $\sigma_2^2$ 已知 | $Z=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$ | $\left(\overline{X}-\overline{Y}\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$ | $(\widehat{\mu_1-\mu_2})_U=\overline{X}-\overline{Y}+z_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$<br>$(\widehat{\mu_1-\mu_2})_L=\overline{X}-\overline{Y}-z_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ |
| | $\mu_1-\mu_2$ | $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知 | $t=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$<br>其中 $S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ | $\left(\overline{X}-\overline{Y}\pm t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$ | $(\widehat{\mu_1-\mu_2})_U=\overline{X}-\overline{Y}+t_\alpha(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$<br>$(\widehat{\mu_1-\mu_2})_L=\overline{X}-\overline{Y}-t_\alpha(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ |
| | $\mu_1-\mu_2$ | $\sigma_1^2\neq\sigma_2^2$ 未知 | $t=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$<br>$\begin{cases}\sim N(0,1), & n_1,n_2>50,\\ \sim t(k), & k=\min(n_1-1,n_2-1)\end{cases}$ | $\left(\overline{X}-\overline{Y}\pm q_{\alpha/2}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)$<br>其中 $q_p=\begin{cases}z_p, & \text{用 } t\sim N(0,1) \text{ 近似},\\ t_p(k), & \text{用 } t\sim t(k) \text{ 近似}\end{cases}$ | $(\widehat{\mu_1-\mu_2})_U=\overline{X}-\overline{Y}+q_\alpha\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$<br>$(\widehat{\mu_1-\mu_2})_L=\overline{X}-\overline{Y}-q_\alpha\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$ |
| | $\frac{\sigma_1^2}{\sigma_2^2}$ | $\mu_1,\mu_2$ 未知 | $F=\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$ | $\left(\frac{S_1^2/S_2^2}{F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2/S_2^2}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right)$ | $\widehat{\left(\frac{\sigma_1^2}{\sigma_2^2}\right)}_U=\frac{S_1^2/S_2^2}{F_{1-\alpha}(n_1-1,n_2-1)}$<br>$\widehat{\left(\frac{\sigma_1^2}{\sigma_2^2}\right)}_L=\frac{S_1^2/S_2^2}{F_{\alpha}(n_1-1,n_2-1)}$ |

第六行不考

# 估计量的评价准则

无偏性准则、有效性准则、均方误差准则、相合性准则
有效性的前提是无偏性（也就是说讨论哪个最有效的前提是这两种估计都是无偏估计）
均方误差准则：MSE

设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\ldots,X_n)$ 是总体参数 $\theta$ 的估计量，称 $E[(\hat{\theta}-\theta)^2]$ 是估计量 $\hat{\theta}$ 的**均方误差**，记为 $Mse(\hat{\theta})$。

- 设 $\widehat{\theta_1},\widehat{\theta_2}$ 都是 $\theta$ 的估计量，若 $\forall\theta\in\Theta, Mse(\widehat{\theta_1})\leq Mse(\widehat{\theta_2})$，则称在均方误差准则下，$\widehat{\theta_1}$ 优于 $\widehat{\theta_2}$
- 均方误差准则常用于**有偏估计**之间，有偏估计与无偏估计之间的比较
- 若 $\hat{\theta}$ 是参数 $\theta$ 的无偏估计量，则 $Mse(\hat{\theta})=Var(\hat{\theta})$，即均方误差准则在**无偏估计**之间的比较等价于有效性准则
- 在实际情况下，均方误差准则比无偏性准则更重要，即如果一个估计量虽然有偏，但其均方误差较小，有时比方差较大的无偏估计更有用

# 第八章 假设检验

考虑假设问题：$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$，取检验统计量 $F=\dfrac{S_1^2}{S_2^2}$。当 $H_0$ 检验时，$F\sim F(n_1-1,n_2-1)$，此时 $F$ 的取值既不能偏大也不能偏小，因此检验的拒绝域为：

$$W=\{F\geq F_{\alpha/2}(n_1-1,n_2-1)\text{ or }F\leq F_{1-\alpha/2}(n_1-1,n_2-1)\}$$

将样本值代入后得到的检验统计量的值记为 $f_0$，即 $f_0=\dfrac{s_1^2}{s_2^2}$，设：

$$p_0=P_{\sigma_1^2=\sigma_2^2}(\frac{S_1^2}{S_2^2}\leq\frac{s_1^2}{s_2^2})=P(F(n_1-1,n_2-1)\leq f_0)$$

P- 值为

- 双侧检验：$2\min\{p_0,1-p_0\}$
- 左侧检验：$p_0$
- 右侧检验：$1-p_0$

| | 原假设 $H_0$ | 检验统计量 | 备择假设 $H_1$ | 拒绝域 | 检验统计量的取值 | $P$－值 |
|---|---|---|---|---|---|---|
| 1 | $\mu \leqslant \mu_0$<br>$\mu \geqslant \mu_0$<br>$\mu = \mu_0$<br>($\sigma^2$ 已知) | $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ | $\mu > \mu_0$<br>$\mu < \mu_0$<br>$\mu \neq \mu_0$ | $Z \geqslant z_\alpha$<br>$Z \leqslant -z_\alpha$<br>$\|Z\| \geqslant z_{\alpha/2}$ | $z_0 = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$ | $1 - \Phi(z_0)$<br>$\Phi(z_0)$<br>$2(1 - \Phi(\|z_0\|))$ |
| 2 | $\mu \leqslant \mu_0$<br>$\mu \geqslant \mu_0$<br>$\mu = \mu_0$<br>($\sigma^2$ 未知) | $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$ | $\mu > \mu_0$<br>$\mu < \mu_0$<br>$\mu \neq \mu_0$ | $T \geqslant t_\alpha(n-1)$<br>$T \leqslant -t_\alpha(n-1)$<br>$\|T\| \geqslant t_{\alpha/2}(n-1)$ | $t_0 = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$ | $P\{t(n-1) \geqslant t_0\}$<br>$P\{t(n-1) \leqslant t_0\}$<br>$2P\{t(n-1) \geqslant \|t_0\|\}$ |
| 3 | $\mu_1 - \mu_2 \leqslant \delta$<br>$\mu_1 - \mu_2 \geqslant \delta$<br>$\mu_1 - \mu_2 = \delta$<br>($\sigma_1^2, \sigma_2^2$ 已知) | $Z = \dfrac{\overline{X} - \overline{Y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ | $\mu_1 - \mu_2 > \delta$<br>$\mu_1 - \mu_2 < \delta$<br>$\mu_1 - \mu_2 \neq \delta$ | $Z \geqslant z_\alpha$<br>$Z \leqslant -z_\alpha$<br>$\|Z\| \geqslant z_{\alpha/2}$ | $z_0 = \dfrac{\overline{x} - \overline{y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ | $1 - \Phi(z_0)$<br>$\Phi(z_0)$<br>$2(1 - \Phi(\|z_0\|))$ |
| 4 | $\mu_1 - \mu_2 \leqslant \delta$<br>$\mu_1 - \mu_2 \geqslant \delta$<br>$\mu_1 - \mu_2 = \delta$<br>($\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知) | $T = \dfrac{\overline{X} - \overline{Y} - \delta}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$<br>$S_w^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$ | $\mu_1 - \mu_2 > \delta$<br>$\mu_1 - \mu_2 < \delta$<br>$\mu_1 - \mu_2 \neq \delta$ | $T \geqslant t_\alpha(n_1 + n_2 - 2)$<br>$T \leqslant -t_\alpha(n_1 + n_2 - 2)$<br>$\|T\| \geqslant t_{\alpha/2}(n_1 + n_2 - 2)$ | $t_0 = \dfrac{\overline{x} - \overline{y} - \delta}{s_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ | $P\{t(n_1 + n_2 - 2) \geqslant t_0\}$<br>$P\{t(n_1 + n_2 - 2) \leqslant t_0\}$<br>$2P\{t(n_1 + n_2 - 2) \geqslant \|t_0\|\}$ |
| 5 | $\mu_1 - \mu_2 \leqslant \delta$<br>$\mu_1 - \mu_2 > \delta$<br>$\mu_1 - \mu_2 = \delta$<br>($\sigma_1^2 \neq \sigma_2^2$ 未知,<br>大样本情形) | $T = \dfrac{\overline{X} - \overline{Y} - \delta}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$ | $\mu_1 - \mu_2 > \delta$<br>$\mu_1 - \mu_2 < \delta$<br>$\mu_1 - \mu_2 \neq \delta$ | $T \geqslant z_\alpha$<br>$T \leqslant -z_\alpha$<br>$\|T\| \geqslant z_{\alpha/2}$ | $t_0 = \dfrac{\overline{x} - \overline{y} - \delta}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ | $1 - \Phi(t_0)$<br>$\Phi(t_0)$<br>$2(1 - \Phi(\|t_0\|))$ |

| | 原假设 $H_0$ | 检验统计量 | 备择假设 $H_1$ | 拒绝域 | 检验统计量的取值 | $P$－值 |
|---|---|---|---|---|---|---|
| 6 | $\mu_1 - \mu_2 \leqslant \delta$<br>$\mu_1 - \mu_2 \geqslant \delta$<br>$\mu_1 - \mu_2 = \delta$<br>($\sigma_1^2 \neq \sigma_2^2$ 未知) | $T = \dfrac{\overline{X} - \overline{Y} - \delta}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$ | $\mu_1 - \mu_2 > \delta$<br>$\mu_1 - \mu_2 < \delta$<br>$\mu_1 - \mu_2 \neq \delta$ | $T \geqslant t_\alpha(k)$<br>$T \leqslant -t_\alpha(k)$<br>$\|T\| \geqslant t_{\alpha/2}(k)$<br>$k = \min\{n_1 - 1, n_2 - 1\}$<br>或采用公式 (8.3.6) | $t_0 = \dfrac{\overline{x} - \overline{y} - \delta}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ | $P\{t(k) \geqslant t_0\}$<br>$P\{t(k) \leqslant t_0\}$<br>$2P\{t(k) \geqslant \|t_0\|\}$ |
| 7 | $\sigma^2 \leqslant \sigma_0^2$<br>$\sigma^2 \geqslant \sigma_0^2$<br>$\sigma^2 = \sigma_0^2$<br>($\mu$ 未知) | $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ | $\sigma^2 > \sigma_0^2$<br>$\sigma^2 < \sigma_0^2$<br>$\sigma^2 \neq \sigma_0^2$ | $\chi^2 \geqslant \chi_\alpha^2(n-1)$<br>$\chi^2 \leqslant \chi_{1-\alpha}^2(n-1)$<br>$\chi^2 \geqslant \chi_{\alpha/2}^2(n-1)$ 或<br>$\chi^2 \leqslant \chi_{1-\alpha/2}^2(n-1)$ | $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$ | $P\{\chi^2(n-1) \geqslant \chi_0^2\}$<br>$P\{\chi^2(n-1) \leqslant \chi_0^2\}$<br>$2\min\{p_0, 1-p_0\}$(其中<br>$p_0 = P\{\chi^2(n-1) \leqslant \chi_0^2\}$) |
| 8 | $\sigma_1^2 \leqslant \sigma_2^2$<br>$\sigma_1^2 \geqslant \sigma_2^2$<br>$\sigma_1^2 = \sigma_2^2$<br>($\mu_1, \mu_2$ 未知) | $F = \dfrac{S_1^2}{S_2^2}$ | $\sigma_1^2 > \sigma_2^2$<br>$\sigma_1^2 < \sigma_2^2$<br>$\sigma_1^2 \neq \sigma_2^2$ | $F \geqslant F_\alpha(n_1 - 1, n_2 - 1)$<br>$F \leqslant F_{1-\alpha}(n_1 - 1, n_2 - 1)$<br>$F \geqslant F_{\alpha/2}(n_1 - 1, n_2 - 1)$ 或<br>$F \leqslant F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$ | $f_0 = \dfrac{s_1^2}{s_2^2}$ | $P\{F(n_1 - 1, n_2 - 1) \geqslant f_0\}$<br>$P\{F(n_1 - 1, n_2 - 1) \leqslant f_0\}$<br>$2\min\{p_0, 1-p_0\}$(其中<br>$p_0 = P\{F(n_1 - 1, n_2 - 1) \leqslant f_0\}$) |
| 9 | $\mu_D \leqslant \delta$<br>$\mu_D \geqslant \delta$<br>$\mu_D = \delta$<br>(成对数据) | $T = \dfrac{\overline{D} - \delta}{S_D/\sqrt{n}}$ | $\mu_D > \delta$<br>$\mu_D < \delta$<br>$\mu_D \neq \delta$ | $T \geqslant t_\alpha(n-1)$<br>$T \leqslant -t_\alpha(n-1)$<br>$\|T\| \geqslant t_{\alpha/2}(n-1)$ | $t_0 = \dfrac{\overline{d} - \delta}{s_d/\sqrt{n}}$ | $P\{t(n-1) \geqslant t_0\}$<br>$P\{t(n-1) \leqslant t_0\}$<br>$2P\{t(n-1) \geqslant \|t_0\|\}$ |

第五行、第六行不考

**两个样本的方差相等且未知，要用到$S_w$**

右边检验P-值取右边

左边检验P-值取左边

双边检验P-值取两边

# 期末历年题

# 2014-2015 P105

三（4）

六（1）

# 2014-2015 P94

无

# 2015-2016 P86

一、2（泊松分布）、4半正态分布$Var(|X|)=\sigma^2(1-\frac{2}{\pi})$（正态分布的概率密度和中心极限定理）

五

# 2015-2016 P78

一、1（事件之间也可以做加减法，事件的加减就是概率的加减）

二、（2）应该明确不同事件之间的关联

# 2016-2017 P71

一、4 求均方误差 可以用$\chi^2$的方差的二级结论
5 二维正态分布

二、注意是均值为3，而不是$\lambda$为3
三、(2)注意边际密度函数要写全(3)
四、(2)要注意的是$Y_1$和$Y_2$，都和X有关，这两者对应的是一个X，也就是说$Y_1$和$Y_2$是不独立的
六、(3)(4)矩估计和极大似然估计
**QAQ这份卷子写的太糟糕了**

# 2016-2017 P64

一、5 这个题后两问没算，有时间的话算一下
六、（2）有效估计量-->方差最小！！！ 终于结束了今天QAQ

# 2017-2018 P55

一、4 均匀分布的均值和方差
四、(3)这种题第一次见，感觉答案的方法有点道理(ˇˍˇ)，有点不太理解
六、(3)计算出错了QAQ

# 2017-2018 P47

这份卷子不难
四、和上一份题的第四题一样

# 2018-2019 P40

一、
4. 做错了QAQ
5. 最后一空挺有意思的，值得再看一下 QAQ今天就到这里了，学不动了QAQ

二、
分布函数用F表示，这题值得重做一遍，易错（要区分分布函数和概率分布率）
F不一定连续，但是F的上限一定是1

# 2020-2021 P30

一、
6 区分单侧置信上限和单侧置信下限和置信区间，分别应该怎么取
二、
只是问什么相关，没有问相关系数是多少
三、
要区分边际密度函数和条件分布下的密度函数，这两者是不一样的

# 2021-2022 P24

什么情况，这么简单？不像历年题

# 2021-2022 P13

一、
3 第二个空要再做一遍，$P(T>t)$就等价于在时间间隔t内不发生故障，也就是求P(0)的概率，得到的是一个关于t的函数
5 置信区间：置信水平为0.95的意思就是真值$\mu$落在这个置信区间的概率为0.95或者说95%
二、
之前一直有个东西理解都是错的，那就是P(XY)=P(X)P(Y)，说明X和Y是独立的，其中的XY并不是X与Y的乘积，而应该写成P(X,Y)，也就是X和Y事件同时发生，比如说P(X=0,Y=0)≠P(X=0)*P(Y=0)，那就说明X和Y是不独立的
三、
第二问写错了，需要再重新写一遍
四、
这个题不用再做了，但是要注意的是第一问算出来$\sigma^2$极大似然估计值应该是含有$\mu$，要把$\mu$的极大似然估计值带进去也就是$\bar{X}$
第二问就是在考极大似然估计值的线性运算
五、
三个重要的抽样分布，$t(n), F(n_1, n_2)$，还有一个打不出来，它们代表的其实都是横轴坐标

# 2022-2023 P5

一、
2(2)泊松分布具有可加性，但是这个题除以10，意味着不是泊松分布，题目应该问的是近似服从什么分布，用中心极限定理
4 题目不难，答案错了
5 第一类错误：弃真，，，第二类错误：存伪，，，第三问P是啥？？？
二、
概率都是非负的
三、
用下面这个二级结论是有条件的：

推广到 $n$ 个变量：设 $X_1, X_2, \ldots, X_n$ 为 $n$ 个相互独立的随机变量，分布函数分别为 $F_1(x), F_2(x), \ldots, F_n(x)$，记 $M = \max\{X_1, X_2, \ldots, X_n\}, N = \min\{X_1, X_2, \ldots, X_n\}$，则：

$$F_M(t) = \prod_{i=1}^{n} F_i(t)$$

$$F_N(t) = 1 - \prod_{i=1}^{n}[1 - F_i(t)]$$

这个结论的前提是相互独立，但是这个题的x和y是不相互独立的，所以不能用这个结论，要从定义入手。
其实从F(m)的表达式就可以看出来，如果按照独立算出来是不连续的，肯定不对

# 2023-2024 p2

没答案QAQ

# 综合题1

一、
1 $P(B-A)$的意思就是B发生而A不发生的概率
五、
用独立这个条件，可以得到$f(x,y) = f(x)f(y)$，进而得到联合概率密度

# 综合题2

一、
1 这个题的第二个空，高中题，可以那样理解：在前六个人里选两个，这是一共的可能的情况，前四个人获奖，分成得一个奖和得两个奖
5 t分布，还得再记记
三、
(3)因为不是独立的，所以要找联合分布

# 综合题3

不难

# 综合题4

一、
3（2）切比雪夫不等式
四、
这个题要注意的是y的范围是(-1,1)，做的很多题都是x、y是正数，但是实际上并不是x、y一定要是正数，这个题也可以算一下面积，算出来会发现，如果x和y都是正数，显然是不成里的
(4)这一问好难QAQ
七、
注意第一问：原假设和备择假设