$X_i-\bar{X}$，$i=1,2,\cdots,n$.

（Ⅰ）求 Y_i 的方差 DY_i，$i=1,2,\cdots,n$；

（Ⅱ）求 Y_1 与 Y_n 的协方差 $\operatorname{cov}(Y_1,Y_n)$；

（Ⅲ）若 $c(Y_1+Y_n)^2$ 是 σ^2 的无偏估计量，求常数 c.

[提示] 本题主要考查随机变量的数字特征、无偏估计．在计算 DY_i 及 $\operatorname{cov}(Y_1,Y_n)$ 时，用 $X_1,\cdots,X_n$ 的线性组合表示 Y_i 之后再利用 X_i 间的独立性，易于计算.

[解]（Ⅰ）
$$\begin{aligned}DY_i&=D(X_i-\bar{X})\\&=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum_{k\neq i}X_k\right]\\&=\frac{n-1}{n}\sigma^2,i=1,2,\cdots,n.\end{aligned}$$

（Ⅱ）
$$\begin{aligned}\operatorname{cov}(Y_1,Y_n)&=E(Y_1-EY_1)(Y_n-EY_n)\\&=E(X_1-\bar{X})(X_n-\bar{X})\\&=E(X_1X_n)+E\bar{X}^2-E(X_1\bar{X})-E(X_n\bar{X})\\&=EX_1EX_n+D\bar{X}-\frac{1}{n}EX_1^2-\frac{1}{n}\sum_{i=2}^{n}E(X_1X_i)\\&\quad-\frac{1}{n}EX_n^2-\frac{1}{n}\sum_{i=1}^{n-1}E(X_iX_n)\\&=-\frac{1}{n}\sigma^2.\end{aligned}$$

（Ⅲ）
$$\begin{aligned}E[c(Y_1+Y_n)^2]&=cD(Y_1+Y_n)\\&=c[DY_1+DY_n+2\operatorname{cov}(Y_1,Y_n)]\\&=c\left(\frac{n-1}{n}+\frac{n-1}{n}-\frac{2}{n}\right)\sigma^2\\&=\frac{2(n-2)}{n}c\sigma^2\\&=\sigma^2,\end{aligned}$$

故
$$c=\frac{n}{2(n-2)}.$$

[典型错误] 误认为 $DY_i=DX_i+D\bar{X}$，$D(Y_1+Y_n)=DY_1+DY_n$.

八、假设检验(数学三)

估计和检验是数理统计中的两大类问题．估计是依据样本推断总体未知参数或分布，检验是依据样本来推断关于总体参数或分布的某种假设是否成立．本节的重点是理解“假设”的概念、类型及显著性检验的思想，掌握显著性检验的基本步骤，会构造简单假设的显著性检验．另外也要理解检验的两类错误，并会计算某些简单情形的两类错误的概率．对于单个及两个正态总体的均值和方差的假设会进行显著性检验.

·考试内容与要求·

考试内容

显著性检验　假设检验的两类错误　单个及两个正态总体的均值和方差的假设检验

考试要求

1．理解“假设”的概念和基本类型：理解显著性检验的基本思想．掌握假设检验的基本步骤：会构造简单假设的显著性检验.

2．理解假设检验可能产生的两类错误，对于较简单的情形，会计算两类错误的概率.

3．掌握单个及两个正态总体的均值和方差的假设检验.

·考试内容解析·

（一）假设检验的基本概念

1. 假设

关于总体的论断或命题、猜测或推测等就是假设．两个二者必居其一的假设 H_0 和 H_1，其中一个称为原假设或零假设，常以 H_0 表示，另一个则称为备选假设或对立假设，常以 H_1 表示．对于原假设的选取，要考虑到易于统计分析．如 H_0：两班成绩相同 $\leftrightarrow H_1$：两班成绩不同．

不论是原假设还是备选假设，若其中只含一个参数值，则称为简单假设，否则称为复合假设．如 H：$p=p_0$ 即为简单假设．而 H：$p\neq p_0$ 或 H：$p<p_0$ 则是复合假设．

2. 假设检验

假设检验就是根据样本，按照某种法则，确定在原假设 H_0 与备选假设 H_1 之中接受其一．

3. 两类错误

当原假设 H_0 为真时，我们却拒绝了 H_0，即认为备选假设 H_1 是正确的，则称为第一类错误．当 H_0 不正确时，我们却接受了 H_0，即认为 H_0 是正确的，则称为第二类错误．

4. 显著性检验

对于一个假设检验法则，当样本容量取定后，我们无法同时使其犯第一类与第二类错误的概率都很小，在此情况下，我们总是控制其犯第一类错误的概率，使其不大于给定的 $\alpha(0<\alpha<1)$．这种检验问题就称为显著性检验问题，给定的数 α 称为显著性水平，一般取 α 为 0.1，0.05，0.01 等值．

在进行显著性检验时，常选用一个统计量 T，称为检验统计量；当 T 在某个区域 W 上时，就拒绝 H_0，否则就不能拒绝 H_0．称区域 W 为 H_0 的拒绝域，而拒绝域 W 的选取是通过控制其第一类错误概率而进行的．

一个显著性检验的基本步骤可以概括为：

(1) 根据实际问题，确定原假设 H_0 和备选假设 H_1．

(2) 规定显著性水平 α $(0<\alpha<1)$．

(3) 建立检验准则．即确定检验统计量及 H_0 的拒绝域 W．

(4) 根据样本值判断是否拒绝 H_0．

（二）一个正态总体均值和方差的假设检验

设总体 $X\sim N(\mu,\sigma^2)$，$X_1,\cdots,X_n$ 是来自总体 X 的简单随机样本，$\bar{X}$ 为样本均值，S^2 为样本方差．μ_0，σ_0^2 为已知常数．记

$$U=\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{n}},\quad T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}},$$

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2},\quad \chi_0^2=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu_0)^2.$$

1. 正态总体均值的检验

当总体方差 $\sigma^2=\sigma_0^2$ 时，用下表中的 U 检验；当 σ^2 未知时，用下表中的 t 检验．

假　设 $H_0\leftrightarrow H_1$	H_0 的拒绝域	
	U 检验	t 检验
$\mu=\mu_0\leftrightarrow\mu\neq\mu_0$	$\{\lvert U\rvert\geqslant u_{\frac{\alpha}{2}}\}$	$\{\lvert T\rvert\geqslant t_{\frac{\alpha}{2}}(n-1)\}$
$\mu\leqslant\mu_0\leftrightarrow\mu>\mu_0$	$\{U\geqslant u_\alpha\}$	$\{T\geqslant t_\alpha(n-1)\}$
$\mu\geqslant\mu_0\leftrightarrow\mu<\mu_0$	$\{U\leqslant -u_\alpha\}$	$\{T\leqslant -t_\alpha(n-1)\}$

其中 u_α，$t_\alpha(n-1)$ 分别是正态分布及 $t(n-1)$ 分布的上侧 α 分位数．

2. 正态总体方差的检验

当总体均值 $\mu=\mu_0$ 已知时，用检验统计量 χ_0^2；当 μ 未知时，用检验统计量 χ^2，具体见下表：

假　设 $H_0 \leftrightarrow H_1$	H_0 的拒绝域	
	$\mu=\mu_0$ 已知	μ 未知
$\sigma=\sigma_0^2 \leftrightarrow \sigma^2 \neq \sigma_0^2$	$\{\chi_0^2 \leqslant \chi_{1-\frac{\alpha}{2}}^2(n)\} \cup \{\chi_0^2 \geqslant \chi_{\frac{\alpha}{2}}^2(n)\}$	$\{\chi^2 \leqslant \chi_{1-\frac{\alpha}{2}}^2(n-1)\} \cup \{\chi^2 \geqslant \chi_{\frac{\alpha}{2}}^2(n-1)\}$
$\sigma^2 \leqslant \sigma_0^2 \leftrightarrow \sigma^2 > \sigma_0^2$	$\{\chi_0^2 \geqslant \chi_\alpha^2(n)\}$	$\{\chi^2 \geqslant \chi_\alpha^2(n-1)\}$
$\sigma^2 \geqslant \sigma_0^2 \leftrightarrow \sigma^2 < \sigma_0^2$	$\{\chi_0^2 \leqslant \chi_{1-\alpha}^2(n)\}$	$\{\chi^2 \leqslant \chi_{1-\alpha}^2(n-1)\}$

（三）两个正态总体均值和方差的检验

设总体 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，$X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_n$ 是分别来自总体 X 和 Y 的简单随机样本，且 X 样本与 Y 样本是相互独立的．$\overline{X}$ 和 $\overline{Y}$ 分别为 X 和 Y 样本的均值，S_x^2 和 S_y^2 分别是 X 样本与 Y 样本的方差．定义

$$S_{xy}^2 = \frac{(m-1)S_x^2 + (n-1)S_y^2}{m+n-2},$$

$$U = \frac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}}, \quad T = \frac{\overline{X} - \overline{Y}}{S_{xy}\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}}, \quad F = \frac{S_x^2}{S_y^2}, \quad F_0 = \frac{n\sum\limits_{i=1}^{m}(X_i - \mu_1)^2}{m\sum\limits_{j=1}^{n}(Y_j - \mu_2)^2}.$$

1. 均值差的检验

当 σ_1^2、σ_2^2 均已知时，用 U 检验；当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时，用 t 检验．

假　设 $H_0 \leftrightarrow H_1$	H_0 的拒绝域	
	U 检验	t 检验
$\mu_1 = \mu_2 \leftrightarrow \mu_1 \neq \mu_2$	$\{\lvert U\rvert \geqslant u_{\frac{\alpha}{2}}\}$	$\{\lvert T\rvert \geqslant t_{\frac{\alpha}{2}}(m+n-2)\}$
$\mu_1 \leqslant \mu_2 \leftrightarrow \mu_1 > \mu_2$	$\{U \geqslant u_\alpha\}$	$\{T \geqslant t_\alpha(m+n-2)\}$
$\mu_1 \geqslant \mu_2 \leftrightarrow \mu_1 < \mu_2$	$\{U \leqslant -u_\alpha\}$	$\{T \leqslant -t_\alpha(m+n-2)\}$

2. 方差比的检验

当 μ_1，μ_2 已知时，用检验统计量 F_0；当 μ_1，μ_2 均未知时，用检验统计量 F．

假　设 $H_0 \leftrightarrow H_1$	H_0 的拒绝域	
	μ_1，μ_2 已知	μ_1，μ_2 未知
$\sigma_1^2 = \sigma_2^2 \leftrightarrow \sigma_1^2 \neq \sigma_2^2$	$\{F_0 \leqslant F_{1-\frac{\alpha}{2}}(m,n)\} \cup \{F_0 \geqslant F_{\frac{\alpha}{2}}(m,n)\}$	$\{F \leqslant F_{1-\frac{\alpha}{2}}(m-1,n-1)\} \cup \{F \geqslant F_{\frac{\alpha}{2}}(m-1,n-1)\}$
$\sigma_1^2 \leqslant \sigma_2^2 \leftrightarrow \sigma_1^2 > \sigma_2^2$	$\{F_0 \geqslant F_\alpha(m,n)\}$	$\{F \geqslant F_\alpha(m-1,n-1)\}$
$\sigma_1^2 \geqslant \sigma_2^2 \leftrightarrow \sigma_1^2 < \sigma_2^2$	$\{F_0 \leqslant F_{1-\alpha}(m,n)\}$	$\{F \leqslant F_{1-\alpha}(m-1,n-1)\}$

·例 题 详 解·

例 3.8.1 设某次考试的考生成绩服从正态分布，从中随机地抽取 36 位考生的成绩，算得平均成绩为 66.5 分，标准差为 15 分．问在显著性水平 0.05 下，是否可以认为这次考试全体考生的平均成绩为 70 分？并给出检验过程．

附表：t 分布表

$$P\{t(n) \leqslant t_p(n)\} = p$$

$t_p(n)$ \ p / n	0.95	0.975
35	1.689 6	2.030 1
36	1.688 3	2.028 1

［提示］ 本题主要考查正态总体均值的双边检验．根据抽样分布性质$\frac{\bar{X}-\mu}{S}\sqrt{n}\sim t(n-1)$．因此，当原假设为真时($\mu=70$)，拒绝域为：

$$|T|=\frac{|\bar{X}-70|}{S}\sqrt{n}\geqslant t_{1-\frac{\alpha}{2}}(n-1),$$

将 $\bar{X}$，S，n 代入计算出 $|T|$ 值，查表得出 $t_{1-\frac{\alpha}{2}}(n-1)$，比较 $|T|$ 与 $t_{1-\frac{\alpha}{2}}(n-1)$之大小，得出接受或拒绝原假设的结论．

［解］ 设该次考试的考生成绩为 X，则 $X\sim N(\mu,\sigma^2)$．把从 X 中抽取的容量为 n 的样本均值记为 $\bar{X}$，样本标准差记为 S．本题是在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\ \mu=70;\ H_1:\ \mu\neq 70,$$

拒绝域为

$$|T|=\frac{|\bar{X}-70|}{S}\sqrt{n}\geqslant t_{1-\frac{\alpha}{2}}(n-1).$$

由 $n=36$，$\bar{X}=66.5$，$S=15$，$t_{0.975}(36-1)=2.030\ 1$，算得

$$|T|=\frac{|66.5-70|\sqrt{36}}{15}=1.4<2.030\ 1,$$

所以接受假设 $H_0:\mu=70$，即在显著性水平 0.05 下，可以认为这次考试全体考生的平均成绩为 70 分．

［典型错误］ 分不开单双边检验及 U 检验与 t 检验的区别．