习题二

2.1 盒中有大小相同的三个球，其中两个球的标号为 0，另一个球的标号为 1，有放回地从盒中随机取球 2 次，记(X_1, X_2)为取到球的标号.

（1）写出总体的分布，并求总体的期望和方差；

（2）写出样本(X_1, X_2)的联合分布；

（3）写出样本均值$\overline{X}$的分布，并求$\overline{X}$的期望和方差.

解 （1）

X	0	1
P	$\frac{2}{3}$	$\frac{1}{3}$

$EX = \frac{1}{3}, EX^2 = \frac{1}{3}, DX = \frac{2}{9}$ ；

（2）

$X_1 \backslash X_2$	0	1
0	$\frac{4}{9}$	$\frac{2}{9}$
1	$\frac{2}{9}$	$\frac{1}{9}$

(3)

$\overline{X}$	0	$\frac{1}{2}$	1
P	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

$E(\overline{X}) = \frac{1}{3}, D\overline{X} = \frac{1}{9}$.

2.2 从一批铁钉中随机地抽取 16 枚，测得它们的长度（单位：cm）为

2.14， 2.10， 2.13， 2.15， 2.13， 2.12， 2.13， 2.10，
2.15， 2.12， 2.14， 2.10， 2.13， 2.11， 2.14， 2.11。

（1）求样本均值$\overline{X}$，修正样本方差S^{*2}，修正样本标准差S^*，样本方差S^2和样本标准差S的观测值；

（2）求样本极差R和样本中位数$\mathrm{med}(X_1,\cdots,X_n)$的观测值。

解　(1)用计算器的统计功能可以求得　$\overline{X}=2.125$，$S^{*2}=0.00029333$，$S^*=0.017127$，$S^2=0.000275$，$S=0.016583$；

（2）将样本观测值按照从小到大的次序排列，可以求得

$$R=X_{(n)}-X_{(1)}=X_{(16)}-X_{(1)}=2.15-2.10=0.05 \quad ;$$

$$\mathrm{med}(X_1,\cdots,X_n)=\frac{X_{(\frac{n}{2})}+X_{(\frac{n}{2}+1)}}{2}=\frac{X_{(8)}+X_{(9)}}{2}=\frac{2.13+2.13}{2}=2.13 \quad 。$$

2.3　设　$(X_1,X_2,\cdots,X_n)$，$(Y_1,Y_2,\cdots,Y_n)$　是两个样本，它们之间有下列关系：

$$Y_i=\frac{X_i-a}{b} \quad , \quad i=1,2,\cdots,n \quad ,$$

其中　a，$b\neq 0$　是常数。求：

（1）它们的样本均值　$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$　与　$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$　之间的关系；

（2）它们的样本方差　$S_x^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$　与　$S_y^2=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\overline{Y})^2$　之间的关系。

解　（1）$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i=\frac{1}{n}\sum_{i=1}^{n}\frac{X_i-a}{b}=\frac{\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}na}{b}=\frac{\overline{X}-a}{b}$　；

（2）$S_y^2=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\overline{Y})^2=\frac{1}{n}\sum_{i=1}^{n}(\frac{X_i-a}{b}-\frac{\overline{X}-a}{b})^2=\frac{1}{nb^2}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{S_x^2}{b^2}$　。

2.4　设有样本$(X_1,X_2,\cdots,X_n)$，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$是样本均值，$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$是样本方差，$\mu$是常数，证明

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2=S^2+(\overline{X}-\mu)^2 \quad 。$$

证 $$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\frac{2}{n}\sum_{i=1}^{n}X_i\mu+\frac{1}{n}\sum_{i=1}^{n}\mu^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}^2+\overline{X}^2-2\overline{X}\mu+\mu^2=S^2+(\overline{X}-\mu)^2\ 。$$

2.5 设 $\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ 和 $S_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2$ 分别是样本 $(X_1,X_2,\cdots,X_n)$ 的样本均值和样本方差，现在样本中增加一个新观测值 X_{n+1}，相应地，样本均值和样本方差变为 $\overline{X}_{n+1}=\frac{1}{n+1}\sum_{i=1}^{n+1}X_i$ 和 $S_{n+1}^2=\frac{1}{n+1}\sum_{i=1}^{n+1}(X_i-\overline{X}_{n+1})^2$，证明：

（1）$\overline{X}_{n+1}=\overline{X}_n+\frac{1}{n+1}(X_{n+1}-\overline{X}_n)$； （2）$S_{n+1}^2=\frac{n}{n+1}\left[S_n^2+\frac{1}{n+1}(X_{n+1}-\overline{X}_n)^2\right]$ 。

证 （1） $$\overline{X}_{n+1}=\frac{1}{n+1}\sum_{i=1}^{n+1}X_i=\frac{1}{n+1}\left(\sum_{i=1}^{n}X_i+X_{n+1}\right)=\frac{n}{n+1}\overline{X}_n+\frac{1}{n+1}X_{n+1}$$

$$=(1-\frac{1}{n+1})\overline{X}_n+\frac{1}{n+1}X_{n+1}=\overline{X}_n+\frac{1}{n+1}(X_{n+1}-\overline{X}_n)\ ;$$

（2） $$S_{n+1}^2=\frac{1}{n+1}\sum_{i=1}^{n+1}(X_i-\overline{X}_{n+1})^2=\frac{1}{n+1}\sum_{i=1}^{n+1}X_i^2-\overline{X}_{n+1}^2$$

$$=\frac{1}{n+1}\left(\sum_{i=1}^{n}X_i^2+X_{n+1}^2\right)-\left(\frac{n}{n+1}\overline{X}_n+\frac{1}{n+1}X_{n+1}\right)^2$$

$$=\frac{n}{n+1}\left[\frac{1}{n}\sum_{i=1}^{n}X_i^2+\frac{1}{n}X_{n+1}^2-\frac{n}{n+1}\overline{X}_n^2-\frac{2}{n+1}\overline{X}_nX_{n+1}-\frac{1}{n(n+1)}X_{n+1}^2\right]$$

$$=\frac{n}{n+1}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}_n^2+\frac{1}{n+1}\overline{X}_n^2-\frac{2}{n+1}\overline{X}_nX_{n+1}+\frac{1}{n+1}X_{n+1}^2\right)$$

$$=\frac{n}{n+1}\left[S_n^2+\frac{1}{n+1}(X_{n+1}-\overline{X}_n)^2\right]\ 。$$

2.6 已知总体 ξ 服从指数分布，概率密度为

$$\varphi(x)=\begin{cases}\lambda\,\mathrm{e}^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

其中，参数$\lambda>0$，$(X_1,X_2,\cdots,X_n)$ 是 ξ 的样本，$\overline{X}$ 是样本均值，S^2 是样本方差，S^{*2} 是修正样本方差，求 $E\overline{X}$，$D\overline{X}$，$E(S^2)$ 和 $E(S^{*2})$。

解 因为总体 ξ 服从参数为 λ 的指数分布，所以 $E\xi=\dfrac{1}{\lambda}$，$D\xi=\dfrac{1}{\lambda^2}$ 。

由定理 2.1 可知

$E\overline{X}=E\xi=\dfrac{1}{\lambda}$，$D\overline{X}=\dfrac{D\xi}{n}=\dfrac{1}{n\lambda^2}$，$E(S^2)=\dfrac{n-1}{n}D\xi=\dfrac{n-1}{n\lambda^2}$，$E(S^{*2})=D\xi=\dfrac{1}{\lambda^2}$ 。

2.7 设 $(X_1,X_2,\cdots,X_n)$ 是总体 $\xi\sim N(\mu,\sigma^2)$ 的样本，$(Y_1,Y_2,\cdots,Y_n)$ 是总体 $\eta\sim N(\mu,\sigma^2)$ 的样本，两个样本相互独立，$\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$，$\overline{Y}=\dfrac{1}{n}\sum\limits_{i=1}^{n}Y_i$ 是 ξ，η 的样本均值，求统计量 $\sum\limits_{i=1}^{n}(X_i+Y_i-\overline{X}-\overline{Y})^2$ 的数学期望。

解法一 因为$(X_1,X_2,\cdots,X_n)$是$\xi\sim N(\mu,\sigma^2)$的样本，$(Y_1,Y_2,\cdots,Y_n)$是$\eta\sim N(\mu,\sigma^2)$的样本，两个样本相互独立，所以

$$E(X_i)=E\xi=\mu，E(Y_i)=E\eta=\mu，i=1,2,\cdots,n 。$$

$$E(\overline{X})=E\xi=\mu，E(\overline{Y})=E\eta=\mu 。$$

$$E(S_x^2)=\frac{n-1}{n}D\xi=\frac{n-1}{n}\sigma^2，E(S_y^2)=\frac{n-1}{n}D\eta=\frac{n-1}{n}\sigma^2 。$$

因此有

$$E\left[\sum_{i=1}^{n}(X_i+Y_i-\overline{X}-\overline{Y})^2\right]$$

$$=E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2+2\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})+\sum_{i=1}^{n}(Y_i-\overline{Y})^2\right]$$

$$=E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2\right]+2E\left[\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})\right]+E\left[\sum_{i=1}^{n}(Y_i-\overline{Y})^2\right]$$

$$=E(nS_x^2)+2\sum_{i=1}^{n}[E(X_i)-E(\overline{X})][E(Y_i)-E(\overline{Y})]+E(nS_y^2)$$

$$=nE(S_x^2)+2\sum_{i=1}^{n}(\mu-\mu)(\mu-\mu)+nE(S_y^2)$$

$$=n\cdot\frac{n-1}{n}\sigma^2+0+n\cdot\frac{n-1}{n}\sigma^2=2(n-1)\sigma^2 \quad 。$$

解法二　因为$(X_1,X_2,\cdots,X_n)$是$\xi\sim N(\mu,\sigma^2)$的样本，$(Y_1,Y_2,\cdots,Y_n)$是$\eta\sim N(\mu,\sigma^2)$的样本，两个样本相互独立，所以

$$X_i\sim N(\mu,\sigma^2),\ Y_i\sim N(\mu,\sigma^2),\ i=1,2,\cdots,n，相互独立。$$

令$Z_i=X_i+Y_i$（$i=1,2,\cdots,n$），则有

$$Z_i=X_i+Y_i\sim N(2\mu,2\sigma^2)\ （i=1,2,\cdots,n），而且相互独立。$$

$(Z_1,Z_2,\cdots,Z_n)$可以看作是总体$\zeta=\xi+\eta\sim N(2\mu,2\sigma^2)$的样本，它的样本均值

$$\bar{Z}=\frac{1}{n}\sum_{i=1}^n Z_i=\frac{1}{n}\sum_{i=1}^n(X_i+Y_i)=\frac{1}{n}\sum_{i=1}^n X_i+\frac{1}{n}\sum_{i=1}^n Y_i=\bar{X}+\bar{Y} \quad ，$$

它的样本方差

$$S_z^2=\frac{1}{n}\sum_{i=1}^n(Z_i-\bar{Z})^2=\frac{1}{n}\sum_{i=1}^n(X_i+Y_i-\bar{X}-\bar{Y})^2 \quad 。$$

所以，

$$E\left[\sum_{i=1}^n(X_i+Y_i-\bar{X}-\bar{Y})^2\right]=E(nS_z^2)=nE(S_z^2)=n\cdot\frac{n-1}{n}D\zeta$$

$$=n\cdot\frac{n-1}{n}\cdot 2\sigma^2=2(n-1)\sigma^2 。$$

2.8　设(X_1,X_2,X_3,X_4,X_5)是总体$\xi\sim N(0,1)$的样本。

（1）求常数a,b，使得$a(X_1+X_2)^2+b(X_3+X_4+X_5)^2$服从$\chi^2(2)$分布，并指出其自由度；

（2）求常数c，使得$\dfrac{c(X_1^2+X_2^2)}{(X_3+X_4+X_5)^2}$服从$F$分布，并指出其自由度。

解　（1）因为（$X_1,X_2,\cdots,X_5$）是$\xi\sim N(0,1)$的样本，所以$X_i\sim N(0,1)$，$i=1,2,\cdots,5$，$X_1,X_2,\cdots,X_5$ 相互独立，所以

$$X_1+X_2\sim N(0,2),\ X_3+X_4+X_5\sim N(0,3)\ ，而且相互独立，$$

即有

$$\frac{X_1+X_2}{\sqrt{2}}\sim N(0,1),\ \frac{X_3+X_4+X_5}{\sqrt{3}}\sim N(0,1)，而且相互独立。$$

由 χ^2 分布的定义可知

$$\frac{(X_1+X_2)^2}{2}+\frac{(X_3+X_4+X_5)^2}{3}=\left(\frac{X_1+X_2}{\sqrt{2}}\right)^2+\left(\frac{X_3+X_4+X_5}{\sqrt{3}}\right)^2\sim\chi^2(2)。$$

可见，只有当 $a=\frac{1}{2}$，$b=\frac{1}{3}$ 时，$a(X_1+X_2)^2+b(X_3+X_4+X_5)^2$ 才服从 χ^2 分布，其自由度为 2。

（2） 因为 $X_1\sim N(0,1)$，$X_2\sim N(0,1)$，X_1,X_2 相互独立，所以由 χ^2 分布的定义可知

$$X_1^2+X_2^2\sim\chi^2(2)。$$

又因为 $\frac{X_3+X_4+X_5}{\sqrt{3}}\sim N(0,1)$，所以由 χ^2 分布的定义可知

$$\left(\frac{X_3+X_4+X_5}{\sqrt{3}}\right)^2\sim\chi^2(1)。$$

而且，它与 $X_1^2+X_2^2$ 相互独立（因为 $X_1,X_2,\cdots,X_5$ 相互独立）。

因此，由 F 分布的定义可知，

$$\frac{3(X_1^2+X_2^2)}{2(X_3+X_4+X_5)^2}=\frac{(X_1^2+X_2^2)/2}{\left(\frac{X_3+X_4+X_5}{\sqrt{3}}\right)^2\Big/1}\sim F(2,1)。$$

可见，只有当 $c=\frac{3}{2}$ 时，$\frac{c(X_1^2+X_2^2)}{(X_3+X_4+X_5)^2}$ 才服从 F 分布，其自由度为 $(2,1)$。

2.9 设 $(X_1,X_2,\cdots,X_m)$ 是 $\xi\sim N(0,\sigma^2)$ 的样本，$(Y_1,Y_2,\cdots,Y_n)$ 是 $\eta\sim N(0,\sigma^2)$ 的样本，两个样本相互独立，证明：

（1）$\frac{\sum_{i=1}^m X_i{}^2+\sum_{j=1}^n Y_j^2}{\sigma^2}\sim\chi^2(m+n)$；（2）$\frac{\sum_{i=1}^m X_i}{\sqrt{\sum_{j=1}^n Y_j^2}}\sqrt{\frac{n}{m}}\sim t(n)$。

证 （1）因为 $(X_1,X_2,\cdots,X_m)$ 是 $\xi\sim N(0,\sigma^2)$ 的样本，所以 $X_i\sim N(0,\sigma^2)$，$i=1,2,\cdots,m$，$X_1,X_2,\cdots,X_m$ 相互独立。

即有 $\frac{X_i}{\sigma}\sim N(0,1)$，$i=1,2,\cdots,m$，$\frac{X_1}{\sigma},\frac{X_2}{\sigma},\cdots,\frac{X_m}{\sigma}$ 相互独立。

由 χ^2 分布定义可知 $\dfrac{\sum_{i=1}^{m} X_i^{\ 2}}{\sigma^2} = \sum_{i=1}^{m}\left(\dfrac{X_i}{\sigma}\right)^2 \sim \chi^2(m)$。

同理可证 $\dfrac{\sum_{j=1}^{n} Y_j^{\ 2}}{\sigma^2} \sim \chi^2(n)$ 。

而且由于两个样本相互独立，所以 $\dfrac{\sum_{i=1}^{m} X_i^{\ 2}}{\sigma^2}$ 与 $\dfrac{\sum_{j=1}^{n} Y_j^{\ 2}}{\sigma^2}$ 相互独立。

因此，由 χ^2 分布的可加性（定理 2.3）可知 $\dfrac{\sum_{i=1}^{m} X_i^{\ 2} + \sum_{j=1}^{n} Y_j^2}{\sigma^2} \sim \chi^2(m+n)$

（2） 因为 $(X_1, X_2, \cdots, X_m)$ 是 $\xi \sim N(0, \sigma^2)$ 的样本，所以 $X_i \sim N(0, \sigma^2)$，$i = 1, 2, \cdots, m$ ，而且相互独立。因此有 $\sum_{i=1}^{m} X_i \sim N(\,0, m\sigma^2)$，所以

$$\frac{\sum_{i=1}^{m} X_i - 0}{\sqrt{m\sigma^2}} = \frac{1}{\sigma\sqrt{m}}\sum_{i=1}^{m} X_i \sim N(\,0, 1) \text{ 。}$$

同时，在上面（1）中已经证得 $\dfrac{\sum_{j=1}^{n} Y_j^{\ 2}}{\sigma^2} \sim \chi^2(n)$，而且由于两个样本相互独立，所以 $\dfrac{1}{\sigma\sqrt{m}}\sum_{i=1}^{m} X_i$ 与 $\dfrac{\sum_{j=1}^{n} Y_j^{\ 2}}{\sigma^2}$ 相互独立。

因此，由 t 分布的定义可知

$$\frac{\sum_{i=1}^{m} X_i}{\sqrt{\sum_{j=1}^{n} Y_j^2}}\sqrt{\frac{n}{m}} = \frac{\dfrac{1}{\sigma\sqrt{m}}\sum_{i=1}^{m} X_i}{\sqrt{\dfrac{\sum_{j=1}^{n} Y_j^{\ 2}}{\sigma^2}\Big/ n}} \sim t(n) \text{ 。}$$

2.10 证明：若 $T \sim t(n)$，则 $T^2 \sim F(1, n)$ 。

证 因为 $T \sim t(n)$，由 t 分布定义可知，必有 $\xi \sim N(0,1)$，$\eta \sim \chi^2(n)$，两者相互独立，

使得 $T=\dfrac{\xi}{\sqrt{\eta/n}}$ ，这时 $T^2=\left(\dfrac{\xi}{\sqrt{\eta/n}}\right)^2=\dfrac{\xi^2/1}{\eta/n}$ 。

因为 $\xi \sim N(0,1)$，由χ^2分布定义可知 $\xi^2 \sim \chi^2(1)$，而且因为 ξ 与 η 相互独立，所以 ξ^2 与 $\eta \sim \chi^2(n)$ 相互独立，因此，由F分布定义可知

$$T^2=\frac{\xi^2/1}{\eta/n} \sim F(1,n) \text{ 。}$$

2.11 设 ξ 服从参数为 λ 的指数分布 $\xi \sim E(\lambda)$，试证明 ξ 的左侧 p 分位数 $E_p(\lambda)=-\dfrac{1}{\lambda}\ln(1-p)$.

解 由题意知，$\xi \sim E(\lambda)$，$F(x)=\begin{cases}1-e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$ ，

$p=F(E_p(\lambda))=1-e^{-\lambda E_p(\lambda)}$得$E_p(\lambda)=-\dfrac{1}{\lambda}\ln(1-p)$.

2.12 设（$X_1,X_2,\cdots,X_m,X_{m+1},\cdots,X_{m+n}$）是总体 $\xi \sim N(\mu,\sigma^2)$ 的样本，证明：

$$\frac{\frac{1}{m}\sum_{i=1}^{m}(X_i-\mu)^2}{\frac{1}{n}\sum_{i=m+1}^{m+n}(X_i-\mu)^2} \sim F(m,n) \text{ 。}$$

证 因为（$X_1,X_2,\cdots,X_m,X_{m+1},\cdots,X_{m+n}$）是总体 $\xi \sim N(\mu,\sigma^2)$ 的样本，所以

$$X_i \sim N(\mu,\sigma^2)\text{，}\quad \frac{X_i-\mu}{\sigma} \sim N(0,1)\text{，}\quad i=1,2,\cdots,m+n \text{ ，}$$

而且它们相互独立。

由χ^2分布定义可知

$$\frac{1}{\sigma^2}\sum_{i=1}^{m}(X_i-\mu)^2=\sum_{i=1}^{m}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(m),$$

$$\frac{1}{\sigma^2}\sum_{i=m+1}^{m+n}(X_i-\mu)^2=\sum_{i=m+1}^{m+n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n),$$

而且两者相互独立。

所以，由 F 分布定义可知

$$\frac{\frac{1}{m}\sum_{i=1}^{m}(X_i-\mu)^2}{\frac{1}{n}\sum_{i=m+1}^{m+n}(X_i-\mu)^2}=\frac{\frac{1}{\sigma^2}\sum_{i=1}^{m}(X_i-\mu)^2\Big/m}{\frac{1}{\sigma^2}\sum_{i=m+1}^{m+n}(X_i-\mu)^2\Big/n}\sim F(m,n) \text{ 。}$$

2.13 设（$X_1,X_2,\cdots,X_m$）是总体 $\xi\sim N(\mu,\sigma^2)$ 的样本，$\overline{X}$ 是样本均值，S^{*2} 是修正样本方差，另有 $X_{m+1}\sim N(\mu,\sigma^2)$，$X_{m+1}$ 与 $X_1,X_2,\cdots,X_m$ 相互独立，证明：

$$\frac{\overline{X}-X_{m+1}}{S^*}\sqrt{\frac{m}{m+1}}\sim t(m-1) \text{ 。}$$

证法一

因为 $(X_1,X_2,\cdots,X_m)$ 是总体 $\xi\sim N(\mu,\sigma^2)$ 的样本，所以由定理 2.5 可知，

$$\overline{X}\sim N(\mu,\frac{\sigma^2}{m}) \text{。}$$

另外已知 $X_{m+1}\sim N(\mu,\sigma^2)$，它与 $X_1,X_2,\cdots,X_m$ 相互独立，因此它也与 $\overline{X}$ 相互独立，由正态分布的可加性可知 $\overline{X}-X_{m+1}\sim N(0,\frac{\sigma^2}{m}+\sigma^2)$，即有

$$\frac{\overline{X}-X_{m+1}}{\sqrt{\frac{m+1}{m}\sigma^2}}\sim N(0,1) \text{。}$$

由定理 2.8（Fisher 引理）可知，$\frac{(m-1)S^{*2}}{\sigma^2}\sim\chi^2(m-1)$，而且 $\overline{X}$ 与 S^{*2} 相互独立。另外又已知 X_{m+1} 与 $X_1,X_2,\cdots,X_m$ 相互独立，因此 X_{m+1} 也与 S^{*2} 相互独立，所以 $\frac{\overline{X}-X_{m+1}}{\sqrt{\frac{m+1}{m}\sigma^2}}$ 与 $\frac{(m-1)S^{*2}}{\sigma^2}$ 相互独立。

因此，由 t 分布的定义可知

$$\frac{\overline{X}-X_{m+1}}{S^*}\sqrt{\frac{m}{m+1}}=\frac{\dfrac{\overline{X}-X_{m+1}}{\sqrt{\dfrac{m+1}{m}\sigma^2}}}{\sqrt{\dfrac{(m-1)S^{*2}}{\sigma^2}\Big/(m-1)}}\sim t(m-1)\quad 。$$

证法二

因为 $X_{m+1}\sim N(\mu,\sigma^2)$，$X_{m+1}$ 可看作是另一个总体 $\eta\sim N(\mu,\sigma^2)$ 的样本，样本容量 $n=1$，样本均值 $\overline{Y}=X_{m+1}$。

由定理 2.12 可知 $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}\sim t(m+n-2)$ ，其中

$$(\overline{X}-\overline{Y})-(\mu_1-\mu_2)=(\overline{X}-X_{m+1})-(\mu-\mu)=\overline{X}-X_{m+1}\quad ,$$

$$S_w=\sqrt{\frac{(m-1)S_x^{*2}+(n-1)S_y^{*2}}{m+n-2}}=\sqrt{\frac{(m-1)S^{*2}+(1-1)S_y^{*2}}{m+1-2}}=S^*\quad ,$$

$$\sqrt{\frac{1}{m}+\frac{1}{n}}=\sqrt{\frac{1}{m}+\frac{1}{1}}=\sqrt{\frac{m+1}{m}}\quad ,\quad m+n-2=m+1-2=m-1\quad 。$$

所以有

$$\frac{\overline{X}-X_{m+1}}{S^*}\sqrt{\frac{m}{m+1}}=\frac{\overline{X}-X_{m+1}}{S^*\sqrt{\dfrac{m+1}{m}}}=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}\sim t(m-1)\quad 。$$

2.14 设总体 $\xi\sim N(\mu_1,a\sigma^2)$ ，$\eta\sim N(\mu_2,b\sigma^2)$ ，其中 $a>0$，$b>0$ 是已知常数，$a\neq b$ 。$(X_1,X_2,\cdots,X_n)$ 是 ξ 的样本，$(Y_1,Y_2,\cdots,Y_n)$ 是 η 的样本，两个样本相互独立，$\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ ，$\overline{Y}=\dfrac{1}{n}\sum\limits_{j=1}^{n}Y_j$ 是 ξ,η 的样本均值，$S_x^{*2}=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ ，$S_y^{*2}=\dfrac{1}{n-1}\sum\limits_{j=1}^{n}(Y_j-\overline{Y})^2$ 是 ξ,η 的修正样本方差。证明：

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{(\dfrac{S_x^{*2}}{a}+\dfrac{S_y^{*2}}{b})(a+b)}}\sqrt{2n}\ \sim\ t(2n-2)\quad 。$$

证 因为 $\xi\sim N(\mu_1,a\sigma^2)$ ，$\eta\sim N(\mu_2,b\sigma^2)$，所以由定理 2.5 可知

$$\bar{X} \sim N(\mu_1, \frac{a\sigma^2}{n}) \quad , \quad \bar{Y} \sim N(\mu_2, \frac{b\sigma^2}{n}) \quad ,$$

而且它们相互独立（因为两个样本相互独立），因此有

$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{(a+b)\sigma^2}{n}) \quad ,$$

即有

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(a+b)\sigma^2}{n}}} \sim N(0,1) \quad 。$$

同时，又因为 $\xi \sim N(\mu_1, a\sigma^2)$ ，$\eta \sim N(\mu_2, b\sigma^2)$，所以由定理 2.8（Fisher 引理）可知

$$\frac{(n-1)S_x^{*2}}{a\sigma^2} \sim \chi^2(n-1) \quad , \quad \frac{(n-1)S_y^{*2}}{b\sigma^2} \sim \chi^2(n-1) \quad ,$$

而且它们相互独立（因为两个样本相互独立），因此，根据 χ^2 分布的可加性，有

$$\frac{(n-1)S_x^{*2}}{a\sigma^2} + \frac{(n-1)S_y^{*2}}{b\sigma^2} \sim \chi^2(2n-2) \quad 。$$

因为 $\bar{X}$ 与 S_x^{*2} 独立，$\bar{Y}$ 与 S_y^{*2} 独立，两个样本又相互独立，所以

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(a+b)\sigma^2}{n}}} \text{ 与 } \frac{(n-1)S_x^{*2}}{a\sigma^2} + \frac{(n-1)S_y^{*2}}{b\sigma^2} \text{ 相互独立。}$$

因此，由 t 分布的定义可知

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\frac{S_x^{*2}}{a} + \frac{S_y^{*2}}{b})(a+b)}}\sqrt{2n} = \frac{\dfrac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(a+b)\sigma^2}{n}}}}{\sqrt{\left(\dfrac{(n-1)S_x^{*2}}{a\sigma^2} + \dfrac{(n-1)S_y^{*2}}{b\sigma^2}\right)\Big/(2n-2)}} \sim t(2n-2) \quad 。$$

2.15 设总体 $\xi \sim N\left(\mu, 2^2\right)$，$\left(X_1, X_2, \ldots, X_n\right)$为其样本，问样本容量$n$至少为多大时，才能保证$P\left\{\left|\bar{X} - \mu\right| \le 0.1\right\} \ge 0.95$？

解 $\xi \sim N\left(\mu, 2^2\right)$，$\left(X_1, X_2, \ldots, X_n\right)$为$\xi$的样本，由$P\left\{\left|\bar{X} - \mu\right| \le 0.1\right\} \ge 0.95$得

$$P\left\{\left|\frac{\overline{X}-\mu}{2/\sqrt{n}}\right| \le \frac{0.1}{2/\sqrt{n}}\right\} = 2\Phi\left(\frac{\sqrt{n}}{20}\right)-1 \ge 0.95,\ \text{即}\frac{\sqrt{n}}{20} \ge 1.96,\text{得} n \ge 1536.64.$$

2.16 设 ξ 服从自由度为 n 的 χ^2 分布，试证明：当 n 充分大时，对任意 $c > 0$ 有 $P\{\xi \le c\} \approx \Phi\left(\dfrac{c-n}{\sqrt{2n}}\right)$.

解 由 $\xi \sim \chi^2(n)$，得 $E\xi = n, D\xi = 2n,$，令 $\xi = X_1^2 + X_2^2 + \ldots + X_n^2$，其中 $X_1, X_2, \ldots, X_n$ 独立同服从于标准正态分布，则 $X_1^2, X_2^2, \ldots, X_n^2$ 独立同分布服从于 $\chi^2(1)$，由中心极限定理知 ξ 近似于正态分布 $N(n, 2n)$.

$$P(\xi \le c) \approx P\left\{\frac{\xi-n}{\sqrt{2n}} \le \frac{c-n}{\sqrt{2n}}\right\} = \Phi\left(\frac{c-n}{\sqrt{2n}}\right).$$