推论 6.1　设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，则

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1)。$$

证明　根据定理 6.2，$\dfrac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1)$，$(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$，且相互独立。因此，

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} = \frac{\sqrt{n}(\bar{X} - \mu)/\sigma}{\sqrt{(n-1)S/((n-1)\sigma^2)}} \sim t(n-1)。$$

推论 6.2　设 $X_1, X_2, \cdots, X_{n_1}$ 是来自正态总体 $N(\mu_1, \sigma_1^2)$ 的一个样本，$Y_1, Y_2, \cdots, Y_{n_2}$ 为来自正态总体 $N(\mu_2, \sigma_2^2)$ 的一个样本，且两样本相互独立。记 $\bar{X}$ 和 $\bar{Y}$ 分别为它们的样本均值，样本方差分别为

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_i - \bar{X})^2,$$

$$S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(Y_i - \bar{Y})^2。$$

则

$$F = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \sim F(n_1 - 1, n_2 - 1)。$$

证明　根据定理 6.2 及 F 分布的定义直接得到。

推论 6.3　设 $X_1, X_2, \cdots, X_{n_1}$ 是来自正态总体 $N(\mu_1, \sigma^2)$ 的一个样本，$Y_1, Y_2, \cdots, Y_{n_2}$ 为来自正态总体 $N(\mu_2, \sigma^2)$ 的一个样本，且两样本相互独立。则

$$T = \sqrt{\frac{n_1 n_2(n_1 + n_2 - 2)}{n_1 + n_2}} \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}} \sim t(n_1 + n_2 - 2)。$$

证明　由定理 6.2，有

$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\sigma^2),$$

$$[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2]/\sigma^2 \sim \chi^2(n_1 + n_2 - 2),$$

且两者相互独立，再由 t 分布的定义得到结论。

习　题　六

1. 设总体 $X \sim N(\mu, \sigma^2)$，其中 μ 未知，σ^2 已知，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。问：$X_1 - \mu, \bar{X}, \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2, \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2, \min_{1 \leqslant i \leqslant n} X_i$ 中哪些是统计量，哪些不是统计量？

2. 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $U[0,1]$ 的样本，求样本均值 $\bar{X}$ 的期望与方差。

3. 设 $X_1, X_2, \cdots, X_n$ 为来自均值为 μ，方差为 σ^2 的总体的样本，证明：

(1) $S^2=\dfrac{1}{n-1}\left[\sum\limits_{i=1}^{n}X_i^2-n\bar{X}^2\right]$；

(2) $E(S^2)=\sigma^2$。

4. 若 $\chi^2\sim\chi^2(n)$，证明：$E(\chi^2)=n, D(\chi^2)=2n$。

5. 设 $X_1,X_2,\cdots,X_{10}$ 是来自正态总体 $N(0,5)$ 的样本，试求 $P\left(\sum\limits_{i=1}^{10}X_i^2>80\right)$。

6. 设 $p_n(x)$ 为自由度为 n 的 t 分布的密度函数，求证：

$$\lim_{n\to\infty}p_n(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}。$$

7. 设 $X_1,X_2,\cdots,X_{15}$ 是来自正态总体 $N(0,4)$ 的样本，求 $Y=\dfrac{X_1^2+X_2^2+\cdots+X_{10}^2}{2(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)}$ 的分布。

8. 设 $X_1,X_2,\cdots,X_{2n}(n\geqslant 2)$ 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，$\bar{X}=\dfrac{1}{2n}\sum\limits_{i=1}^{2n}X_i$，$Y=\sum\limits_{i=1}^{n}(X_i+X_{n+i}-2\bar{X})^2$，求 $E(Y)$。

9. 设 $X_1,X_2,\cdots,X_{n+1}$ 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，$\bar{X}=\dfrac{1}{n}\sum\limits_{i-1}^{n}X_i$，$S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$，求 $\dfrac{X_{n+1}-\bar{X}}{S}\sqrt{\dfrac{n}{n+1}}$ 的分布。

10. 设 $X_1,X_2,\cdots,X_5$ 是来自正态总体 $N(12,4)$ 的样本，试求：

(1) $\bar{X}$ 大于 13 的概率；

(2) $\min\limits_{1\leqslant i\leqslant 5}X_i$ 小于 10 的概率；

(3) $\max\limits_{1\leqslant i\leqslant 5}X_i$ 大于 15 的概率。

11. 设从同一个总体中抽得两组样本，容量分别为 n_1 与 n_2，样本均值，样本方差分别记为 $\bar{X},\bar{Y},S_1^2,S_2^2$。现将两组样本合并在一起，试将联合样本的均值和方差用 $\bar{X},\bar{Y},S_1^2,S_2^2$ 表示出来。

12. 设 $X_1,X_2,\cdots,X_{n_1}$是来自正态总体 $N(\mu_1,\sigma^2)$ 的样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 是来自正态总体 $N(\mu_2,\sigma^2)$ 的样本，试求 $\dfrac{\sum\limits_{i=1}^{n_1}(X_i-\bar{X})^2+\sum\limits_{j=1}^{n_2}(Y_i-\bar{Y})^2}{n_1+n_2-2}$ 的期望和方差。

于是 $\bar{X} \approx N\left(p, \dfrac{p(1-p)}{n}\right)$。利用枢轴变量法可以求出 p 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - u_{\alpha/2}\frac{\sqrt{p(1-p)}}{\sqrt{n}}, \bar{X} + u_{\alpha/2}\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right)。$$

这里的 p 未知，于是我们用 $\hat{p} = \bar{X}$ 代入上式，最后得到 p 的置信度为 $1-\alpha$ 的置信区间近似为

$$\left(\bar{X} - u_{\alpha/2}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}, \bar{X} + u_{\alpha/2}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}\right)。\tag{7.19}$$

事实上，注意到

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

只需将公式 (7.18) 中的 S 用 S^* 代替就可以得到公式 (7.19)。

习　题　七

1. 设总体 $X \sim B(n,p)$，其中 n 为已知常数，p 为未知参数。$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。试求 p 的矩估计和极大似然估计。

2. 设总体 X 的密度函数为

$$p(x;\alpha) = \begin{cases} (\alpha+1)x^{\alpha} & 0 \leqslant x \leqslant 1 \\ 0 & 其他 \end{cases},$$

其中 $\alpha > -1$ 为未知参数。$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。试求参数 α 的矩估计和极大似然估计。

3. 设总体 X 服从对数正态分布，其密度函数为

$$p(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}}x^{-1}\exp\{-\frac{1}{2\sigma^2}(\ln x - \mu)^2\}, \quad x > 0$$

$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。试求参数 μ, σ^2 的矩估计和极大似然估计。

4. 设总体 X 的密度函数为

$$p(x;\mu,\theta) = \begin{cases} \dfrac{1}{\theta}e^{-(x-\mu)/\theta} & x \geqslant \mu \\ 0 & 其他 \end{cases},$$

其中 $\theta > 0$，μ, θ 为未知参数。$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。试求参数 μ, θ 的矩估计和极大似然估计。

5. 设总体为 $\left[\theta - \dfrac{1}{2}, \theta + \dfrac{1}{2}\right]$ 上的均匀分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。试求参数 θ 的矩估计和极大似然估计。

6. 设总体为 $[\theta, 2\theta]$ 上的均匀分布，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本。试求参数 θ 的矩估计和极大似然估计。

7. 设总体 X 的分布律为

$$X \sim \begin{bmatrix} 0 & 1 & 2 & 3 \\ \theta^2 & 2\theta(1-\theta) & \theta^2 & 1-2\theta \end{bmatrix}$$

其中 $0<\theta<\dfrac{1}{2}$ 为未知参数。现有 8 个样本观察值为 $3,1,3,0,3,1,2,3$。试求参数 θ 的矩估计和极大似然估计。

8. 设总体 $X\sim P(\lambda)$，其中 λ 为未知参数。$X_1,X_2,\cdots,X_n$ 是来自总体 X 的一个样本。试求 $P(X=0)$ 的极大似然估计。

9. 设总体 $X\sim N(\mu,1)$，其中 μ 为未知参数。X_1,X_2 是来自总体 X 的一组样本。试验证下列两个估计量：

(1) $\hat{\mu}_1=\dfrac{1}{4}X_1+\dfrac{3}{4}X_2$；　　(2) $\hat{\mu}_2=\dfrac{1}{2}X_1+\dfrac{1}{2}X_2$

均为 μ 的无偏估计，并比较它们的均方误差。

10. 设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 均为 θ 的无偏估计，且相互独立。$D(\hat{\theta}_1)=\sigma_1^2, D(\hat{\theta}_2)=\sigma_2^2$，$\sigma_1^2,\sigma_2^2$ 为已知参数。试求常数 c，使得 $\hat{\theta}=c\hat{\theta}_1+(1-c)\hat{\theta}_2$ 的方差达到最小。

11. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $N(\mu,\sigma^2)$ 的一个样本。试选择适当常数 c，使得 $c\sum\limits_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计。

12. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $N(\mu,\sigma^2)$ 的一个样本，其中 μ 已知。试求 σ^2 的极大似然估计，并与样本方差 S^2 比较它们的均方误差。

13. 对于方差 σ^2 已知的正态总体，问需抽取容量 n 为多大的样本，才能使总体均值 μ 的置信度为 $1-\alpha$ 的置信区间的长度不大于 L?

14. 设某种油漆的 9 个样本，其干燥时间 (单位：时) 为

6.0，　5.7，　5.8，　6.5，　7.0，　6.3，　5.6，　6.1，　5.0

假定干燥时间服从正态总体 $N(\mu,\sigma^2)$，其中 σ^2 未知。求 μ 的置信度为 95%的置信区间。

15. 测量钢筋的直径 X，得到一组数据 (单位：厘米) 如下：

5.35，　5.41，　5.32，　5.33，　5.44，　5.36，　5.37，　5.38

设 X 服从正态分布，求 σ 的置信度为 90%的置信区间。

16. 甲、乙两组生产同种导线，现从甲组中随机抽取 4 根，从乙组中随机抽取 5 根，它们的电阻 (单位：欧姆) 分别为

甲组：　0.143，　0.142，　0.143，　0.137

乙组：　0.140，　0.142，　0.136，　0.138，　0.140

假设两组电阻值均服从正态分布，且方差相同。求 $\mu_1-\mu_2$ 的置信度为 95%的置信区间。

17. 设机器 A 和机器 B 生产钢管，随机抽取机器 A 生产的钢管 18 根，测得钢管内径的样本方差为 $S_{\mathrm{A}}^2=0.34(\mathrm{mm}^2)$。抽取机器 B 生产的钢管 13 根，测得钢管内径的样本方差为 $S_{\mathrm{B}}^2=0.29(\mathrm{mm}^2)$。假设总体均为正态分布，求方差比 $\dfrac{\sigma_{\mathrm{A}}^2}{\sigma_{\mathrm{B}}^2}$ 的置信度为 90%的置信区间。

18. 某市随机抽取 1000 个家庭，调查知道其中有 228 家拥有彩色电视机。试求该市拥有彩色电视机家庭比例 p 的置信度为 95%的置信区间。

19. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X\sim P(\lambda)$ 的一个样本，试求 λ 的置信度为 $1-\alpha$ 的近似置信区间。

将其代入 χ^2 检验统计量

$$\chi^2=\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij}-n\hat{p}_{i\cdot}\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\hat{p}_{\cdot j}}=\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij}-n_{i\cdot}n_{\cdot j}/n)^2}{n_{i\cdot}n_{\cdot j}/n}。$$

当原假设成立时，它的渐进分布为自由度 $rs-(r+s-2)-1=(r-1)(s-1)$ 的 χ^2 分布。于是，在给定显著性水平 α 下，检验的拒绝域为 $W=\{\chi^2\geqslant\chi^2_\alpha((r-1)(s-1))\}$。

例 8.16　我们调查了 520 人研究患高血压与冠心病的关系。数据如下表所示。

表 8-6

	患高血压	无高血压	总计
患冠心病	48	36	84
无冠心病	88	348	436
总计	136	384	520

试问：在显著性水平 $\alpha=0.05$ 下，高血压与冠心病有无关系。

解　因为 $n_{11}=48,n_{12}=36,n_{21}=88,n_{22}=348$，又 $n_{1\cdot}=84,n_{2\cdot}=436$，$n_{\cdot1}=136,n_{\cdot2}=384$。于是 χ^2 检验统计量为

$$\begin{aligned}\chi^2=&\frac{(48-84\times136/520)^2}{84\times136/520}+\frac{(36-84\times384/520)^2}{84\times384/520}\\&+\frac{(88-436\times136/520)^2}{436\times136/520}+\frac{(348-436\times384/520)^2}{436\times384/520}=49.8。\end{aligned}$$

查分布表得 $\chi^2_{0.05}(1)=3.84<\chi^2=49.8$。因此，拒绝原假设，即认为患高血压与冠心病有关。

习　题　八

1. 设总体 $X\sim N(\mu,\sigma^2)$，其中 μ,σ^2 为未知常数。试问下面的假设哪些是简单假设，哪些是复合假设：

(1) $H_0:\mu=0,\sigma=1$;　(2) $H_0:\mu=0,\sigma>1$;

(3) $H_0:\mu<0,\sigma=1$;　(4) $H_0:0<\mu<3$.

2. 某公司的桶装油的标准质量为 10 公斤。现随机抽取 10 桶，称得重量为 $10.2,9.7,10.1,10.3,10.1,9.8,9.9,10.4,10.3,9.8$(公斤)。假定每桶油的重量服从正态分布。试问：在显著性水平 $\alpha=0.01$ 下，该公司的桶装油重量是否为 10 公斤。

3. 某批矿砂的 5 个样本的锂含量为 (%)：3.25,3.27,3.24,3.26,3.24。假定总体服从正态分布，试问在显著性水平 $\alpha=0.01$ 下，能否认为这批矿砂的锂含量的均值为 3.25%。

4. 某厂生产的钢索的断裂强度服从正态分布 $N(\mu,40^2)$。现从这批钢索中抽取 9 根测得其断裂强度的平均值 $\bar{x}$，较正常生产的均值 μ 大 20(公斤/厘米 ^{2})。试问：在显著性水平 $\alpha=0.01$ 下，能否认为这批钢索的质量有所提高？

5. 两台车床生产同一类产品。假定产品的某种性能分别服从正态分布 $N(\mu_i,\sigma_i^2)$，i=1,2。且已知 $\sigma_1^2=8,\sigma_2^2=6$，两个样本值分别为 $(26,25,23,26,25,29,21)$，$(30,26,32,28,28,30,29)$。试问在显著性水平 $\alpha=0.05$ 下，两台车床生产的产品质量有无显著差异？

6. 比较 A, B 两种小麦品种的蛋白质含量。随机抽取 A 种小麦 10 个样本，测得 $\bar{X}=14.3, S_1^2=1.62$。随机抽取 B 种小麦 5 个样本，测得 $\bar{Y}=11.7, S_2^2=0.14$。假定这两种小麦蛋白质含量均服从正态分布，且方差相同。试问：在显著性水平 $\alpha=0.01$ 下，两种小麦的蛋白质含量有无差异？

7. 设 $X_1, X_2, \cdots, X_{n_1}$ 是来自正态总体 $N(\mu_1, \sigma^2)$ 的一个样本，$Y_1, Y_2, \cdots, Y_{n_2}$ 为来自正态总体 $N(\mu_2, \sigma^2)$ 的一个样本，且两样本相互独立，参数均未知。试构造假设 $H_0: \mu_1=C\mu_2$ 的双边检验，其中 C 为非零已知常数。显著性水平为 α。

8. 设甲，乙两个试验员，对同样的试验进行分析，各人的分析数据结果如下：

试验号数	1	2	3	4	5	6	7	8
甲	4.3	3.2	3.8	3.5	3.5	4.8	3.3	3.9
乙	3.7	4.1	3.8	3.8	4.6	3.9	2.8	4.4

试问：甲乙两人的试验分析有无显著差异 $(\alpha=0.05)$？

9. 某厂生产的瓶装矿泉水要求标准差 $\sigma=0.02$。随机抽取 20 瓶矿泉水，发现其样本标准差 $S=0.03$。假定瓶装矿泉水重量服从正态分布。试问：在显著性水平 $\alpha=0.05$ 下，能否认为它们达到了标准差 $\sigma=0.02$ 的要求？

10. 两台机床加工同一种零件，测得直径为

机床甲： 20.5 19.8 19.7 20.4 20.1 20.0 19.0 19.9

机床乙： 19.7 20.8 20.5 19.8 19.4 20.6 19.2

假定产品直径均服从正态分布。试问甲乙两台机床加工的精度有无显著差异 $(\alpha=0.05)$？

11. 在豌豆实验中，观测到黄色和绿色豆子的数目分别为 70 和 27，试问：在显著性水平 $\alpha=0.05$ 下，检验黄色和绿色豆子的数目为 3:1 是否正确。

12. 在孟德尔豌豆试验中，同时考虑豌豆的颜色与形状，共有四种组合：(黄，圆)，(黄，非圆)，(绿，圆)，(绿，非圆)。按照孟德尔理论这四类比例为 9:3:3:1。在一次观察中，发现这四类观测到的数目分别为 315，101，108，32。试在显著性水平 $\alpha=0.05$ 下，检验孟德尔的理论是否正确。

13. 为检验一个骰子的均匀性，投掷 60 次，观测到出现 $1,2,3,4,5,6$ 点的次数为 $13,19,11,8,5,4$。试检验这个骰子是均匀的 $(\alpha=0.05)$。

14. 某修理厂希望知道每天送来修理的汽车数是否服从泊松分布。下表给出了该厂 250 天的送修车数：

送修车数	0	1	2	3	4	5	6	7	8	9	10
天数	2	8	21	31	44	48	39	22	17	13	5

试在显著性水平 $\alpha=0.05$ 下，检验原假设 H_0：每天送来修理的车数服从泊松分布。

15. 对某汽车零件制造厂生产的汽缸螺栓口径抽样 100 次，数据如下：

口径区间	实际频数	口径区间	实际频数
10.93～10.95	5	11.01～11.03	17
10.95～10.97	8	11.03～11.05	6
10.97～10.99	20	11.05～11.07	6
10.99～11.01	34	11.07～11.09	4

试检验螺栓口径是否服从正态分布 $(\alpha=0.05)$？

16. 调查某市郊区桑场采桑员和辅助人员桑毛虫皮炎发病情况，数据如下：

	采桑	不采桑	总计
患病	18	12	30
不患病	4	78	82
总计	22	90	112

试问发生皮炎病是否与工种有关 ($\alpha = 0.05$)?

17. 某工厂三种配方商场的产品质量如下表：

	配方 1	配方 2	配方 3	总计
合格品	63	47	65	175
不合格品	16	7	3	26
总计	79	54	68	201

试在显著性水平 $\alpha = 0.05$ 下，检验原假设 H_0：三种配方生产的产品质量无显著差异?