习题参考答案

习题 1.1

1. 该市所有的成年男子（18 岁以上）为总体，被调查的 50×100 名成年男子为样本。
2. 设有 x 条鱼，则涂上红漆的鱼的比例为$\frac{n}{x}$，即从池塘中打捞出带有红漆的鱼的概率为$p=\frac{n}{x}$，从而由$\frac{n}{x}=p=\frac{k}{m}$ 可知该池塘最有可能的鱼数是$x=\frac{nm}{k}$条。
3. 略。
4. 样本非随机。
5. 娱乐场所的人群和公共汽车站的人群都是具有某种特定性质的人群，不能作为全部市民/家庭的简单随机样本。
6. 可以对 2 356 名新生进行简单随机抽样选取 100 人，也可以对男、女学生进行分层抽样。

习题 1.2

1. 不含任何未知参数的样本函数称为统计量，故 T_1，T_4，T_5 和 T_6 是统计量。
2. $\bar{x}=153.8$，$s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}=8.1$。
3. 各区间的组中值分别为 43，53，63，73，83

   故 $\bar{x}=63.39$，$s=8.02$。
4. （1）证明：易得 $\bar{x}_{n+1}=\frac{n\bar{x}+x_{n+1}}{n+1}$

$$s_{n+1}^2=\frac{1}{n}\sum_{i=1}^{n+1}(x_i-\bar{x}_{n+1})^2=\frac{1}{n}\sum_{i=1}^{n+1}\left(x_i-\frac{n\bar{x}+x_{n+1}}{n+1}\right)^2$$

$$=\frac{1}{n}\sum_{i=1}^{n+1}\left(x_i-\bar{x}+\frac{\bar{x}-x_{n+1}}{n+1}\right)^2=\frac{n-1}{n}s^2+\frac{1}{n+1}(x_{n+1}-\bar{x})^2$$

（2）代入数据得

$\bar{x}_{n+1}=168.125$

$s_{n+1}^2=11.05^2$

5. 证：易证 $\bar{x}=\dfrac{n_1\bar{x}_1+n_2\bar{x}_2}{n_1+n_2}$，$s^2$ 同第7题。

6. 参照上题结论，$\bar{x}=10.16$，$s=0.0184$。

7. 证明：$s^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{n_i}(x_{ij}-\bar{x})^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{n_i}(x_{ij}-\bar{x}_i+\bar{x}_i-\bar{x})^2$ 简单化简后可得证。

8. 参照第7题结论：$\bar{x}=10.153$，　$s=0.0204$

9. (1) $\because E(x_{i+1}-x_i)^2=E(x_{x+1}^2-2x_ix_{x+i}+x_i^2)=2\sigma^2$

$$\therefore E\left[c\sum_{i=1}^{n-1}(x_{x+i}-x_i)^2\right]=2c\sigma^2(n-1)=\sigma^2 \quad \therefore c=\frac{1}{2(n-1)}$$

(2) 方法同(1)，$c'=[n(n-1)]^{-1}$

(3) 见1.4.3节例1.4.5。

(4) 成立。

10. 反证法，具体略。

11. $$\sum_{i=1}^{n}(x_i-c)^2=\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-c)^2$$
$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+2\sum_{i=1}^{n}(x_i-\bar{x})(\bar{x}-c)+n(\bar{x}-c)^2$$

其中，第二项等于0，第三项大于等于0，所以结论成立。

习题 1.3

1. (1) $N(52,1.05^2)$。(2) 0.8303。(3) $n\geqslant 66$。

2. 0.8584。

3. (1) $N\left(p,\dfrac{p(1-p)}{n}\right)$；(2) $N\left(\lambda,\dfrac{\lambda}{n}\right)$；(3) $N\left(\dfrac{a+b}{2},\dfrac{(b-a)^2}{12n}\right)$；(4) $N\left(\dfrac{1}{\lambda},\dfrac{1}{n\lambda^2}\right)$。

4. (1) $N(20,0.5^2)$；(2) 0.9546。

5. 由$\dfrac{(n-1)s^2}{\sigma^2}\sim\chi^2(n-1)$，可得$E(s^2)=\sigma^2$，$\mathrm{Var}(s^2)=2\sigma^4/(n-1)$。

6. $p(t)=0.375(1+t^2/4)^{-2.5}$，$-\infty<t<\infty$

$p(0)=0.375$，$E(t)=0$，$\mathrm{Var}(t)=2$

7. $\bar{X}\sim N(100,\dfrac{4}{15})$，$\bar{Y}\sim N(100,\dfrac{1}{5})$，$\bar{X}-\bar{Y}\sim N(0,\dfrac{7}{15})$

故为0.4642。

8. 由F分布的性质可知

$$F\sim F(n_1,n_2)\Rightarrow\frac{1}{F}\sim F(n_2,n_1)$$

$$\therefore\quad X\sim F(n,n)\Rightarrow\frac{1}{X}\sim F(n,n)$$

即 X 与$\frac{1}{X}$同分布。

则有　　$P(X<1)=p(\frac{1}{X}<1)$

即　　$P(X<1)=P(X>1)$

∴　　$P(X<1)=0.5$

9. x_1，x_2 是来自正态总体 $N(0,\sigma^2)$ 的样本，x_1 与 x_2 独立。

∵$\mathrm{Cov}(X,Y-Z)=\mathrm{Cov}(X,Y)-\mathrm{Cov}(X,Z)$

∴$\mathrm{Cov}(x_1+x_2,x_1-x_2)=\mathrm{Cov}(x_1+x_2,x_1)-\mathrm{Cov}(x_1+x_2,x_2)$

即 x_1+x_2 与 x_1-x_2 不相关。

∵$x_1+x_2\sim N(0,2\sigma^2)$，$x_1-x_2\sim N(0,2\sigma^2)$

∴根据正态分布性质，x_1+x_2 与 x_1-x_2 独立。

∴$\frac{(x_1+x_2)^2}{(x_1-x_2)^2}=\frac{(x_1+x_2)^2/2\sigma^2}{(x_1-x_2)^2/2\sigma^2}\sim F(1,1)$ $\left(\frac{(x_1+x_2)^2}{2\sigma^2}\sim\chi^2(1)\right)$

10. ∵$P\left(\frac{(x_1+x_2)^2}{(x_1-x_2)^2+(x_1+x_2)^2}>k\right)=0.05$

∴$P\left(1+\frac{(x_1-x_2)^2}{(x_1+x_2)^2}<\frac{1}{k}\right)=0.05$，即 $P\left(\frac{(x_1-x_2)^2}{(x_1+x_2)^2}<\frac{1}{k}-1\right)=0.05$

由上题结论知$\frac{(x_1-x_2)^2}{(x_1+x_2)^2}\sim F(1,1)$

∴$\frac{1}{k}-1=F_{0.05}(1,1)$

查表可知　$F_{0.95}(1,1)=161.45$

即　$\frac{1}{k}-1=\frac{1}{161.45}\Rightarrow k=0.994$

11. $C=\sqrt{\frac{n}{(n+1)(n-1)}}$，$n=1$。

12. $\frac{s_1^2/(n_1-1)}{s_2^2/(n_2-1)}\sim F(n_1-1,n_2-1)$，$p=0.0798$。

13. 正态总体下$\frac{\sqrt{n}(\bar{x}-\mu)}{s}\sim t(n-1)$，故 $k=-0.4234$。

14. 见 3.12 节例 3.14。

习题 1.4

1. (1)

$x_{(1)}$	0	1	2	3
P	37/64	19/64	7/64	1/64

$x_{(3)}$	0	1	2	3
P	1/64	7/64	19/64	37/64

(2)

$x_{(1)}$ \ $x_{(3)}$	0	1	2	3
0	1/64	6/64	12/64	18/64
1	0	1/64	6/64	12/64
2	0	0	1/64	6/64
3	0	0	0	1/64

(3) 由上表易知不独立。

2. $n=5$，$p(x)=3x^2$，易知$x_{(1)}\sim p_1(x)=15x^2(1-x^3)^4$，$0<x<1$

$$x_5\sim p_5(x)=15x^{15}，0<x<1$$

3. 求次序统计量 $x_{(1)}$ 的分布函数，形状参数为 m，尺度参数为 $\eta/\sqrt[m]{n}$。

4. (1) $P(x_{(1)}>800)=0.000\ 746\ 6$；(2) $P(x_{(n)}<3\ 000)=0.935\ 2$。

5. $R=9$，$\hat{\sigma}_R=2.924$。

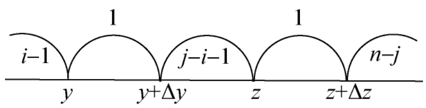
6. $x_{(1)}=29$，$Q_1=36$，$m_d=40$，$Q_3=45.25$，$x_{(50)}=49$，图略。

7. 略。

8. 略。

9. $\bar{x}_\alpha=9.71$。

10. (1) 对于求 ($x_{(i)}$，$x_{(j)}$) ($i<j$) 的联合分布密度函数，可以把数轴进行如下划分：



$x_{(i)}\in(y，y+\Delta y]$，　$x_{(j)}\in(z，z+\Delta z]$

即有 $i-1$ 个观测值小于等于 y，一个落入区间 (y，$y+\Delta y$)，$j-i-1$ 个落入区间 ($y+\Delta y$，z]，一个落入区间 (z，$z+\Delta z$]，余下 $n-j$ 个大于 $z+\Delta z$。

$$\therefore P(x_{(i)}\in(y,y+\Delta y),x_{(j)}\in(z,z+\Delta z))$$

$$=\frac{n!}{(i-1)!\ 1!\ (j-i-1)!\ (n-j)!}[F(y)]^{i-1}p(y)\Delta y[F(z)-F(y+\Delta y)]^{j-i-1}p(z)\Delta z[1-F(z+\Delta z)]$$

$F(x)$ 连续，当 $\Delta y\to 0$，$\Delta z\to 0$ 时，有 $F(y+\Delta y)\to F(y)$，$F(z+\Delta z)\to F(z)$

$$\therefore p_{ij}(y,z)=\lim_{\substack{\Delta y\to 0\\ \Delta z\to 0}}\frac{P(x_{(i)}\in(y+\Delta y),x_{(j)}\in(z,z+\Delta z))}{\Delta y\cdot\Delta z}$$

$$=\frac{n!}{(i-1)!\ (j-i-1)!\ (n-j)!}[F(y)]^{i-1}[F(z)-F(y)]^{j-i-1}[1-F(z)]^{n-j}p(y)p(z)$$

当 $i=1$，$j=n$ 时，即

$$p(u_1,u_2)=n(n-1)[F(u_2)-F(u_1)]^{n-2}p(u_1)p(u_2),u_1<u_2$$

(2) $R=x_{(n)}-x_{(1)}$

$$F(r)=P(x_{(n)}-x_{(1)}<r)=\int_{-\infty}^{+\infty}\int_{u_1}^{u_1+r}p(u_1,u_n)\mathrm{d}u_1\mathrm{d}u_2$$

$$\int_{u_1}^{u_1+r}p(u_1,u_2)\mathrm{d}u_2=\int_{u_1}^{u_1+r}n(n-1)p(u_1)[F(u_2)-F(u_1)]^{n-2}p(u_2)\mathrm{d}u_2$$

$$=np(u_1)[F(u_2)-F(u_1)]^{n-1}\Big|_{u_1}^{u_1+r}$$

$$=np(u_1)[F(u_1+r)-F(u_1)]^{n-1}$$

$\therefore R=x_{(n)}-x_{(1)}$ 的分布函数为：

$$F(r)=n\int_{-\infty}^{+\infty}[F(r+u)-F(u)]^{n-1}p(u)\mathrm{d}u$$

（3）若 $F(x)$ 为均匀分布，则（$x_{(1)}$，$x_{(n)}$）的联合密度函数为：

$p_{1,n}(y,z)=n(n-1)(z-y)^{n-2}$，$0<y<z<1$

由 $R=x_{(n)}-x_{(1)},R>0$ 知 $0<x_{(1)}=x_{(n)}-R\leqslant 1-R$

$$\Rightarrow P(r)=\int_0^{1-r}n(n-1)[(y+r)-y]^{n-2}\mathrm{d}y$$
$$=n(n-1)r^{n-2}(1-r)$$

故为 $Be(n-1,2)$。

样本来自均匀分布 $U(0,\theta)$。

11. 因为 $x_{(1)}$，$x_{(2)}$，…，$x_{(n)}$ 的联合密度函数为：

$$p(x_{(1)},\cdots,x_{(n)})=\frac{n!}{\theta^n}$$

现 $y_1=\frac{x_{(1)}}{x_{(2)}}$，$y_2=\frac{x_{(2)}}{x_{(3)}}$，…，$y_n=x_{(n)}$，故

$x_{(1)}=y_1\cdot y_2\cdots y_n$ $x_{(2)}=y_2\cdots y_n$，…，$x_{(n)}=y_n$

$$|J|=\begin{vmatrix} y_2\cdots y_n & y_1y_3\cdots y_n & \cdots & y_1\cdots y_{n-1}\\ 0 & y_3\cdots y_n & \cdots & y_2\cdots y_{n-1}\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{vmatrix}=y_2\cdot y_3^2\cdot y_4^3\cdots y_n^{n-1}$$

故 y_1，y_2，…，y_n 的联合密度函数为：

$$p(y_1,\cdots,y_n)=\frac{y_2\cdot y_3^2\cdots y_n^{n-1}}{\theta^n}n!$$

而 y_n 的密度函数为 $f(y_n)=\frac{n}{\theta^n}y_n^{n-1}$，$y_{n-1}$ 的密度函数为 $f(y_{n-1})=(n-1)y_{n-1}^{n-2}$，$y_i$ 的密度函数为 $f(y_i)=iy_i^{i-1}$（$i<n$）

$\therefore p(y_1\cdots y_n)=p(y_1)\cdots p(y_n)$

即相互独立。

12. x_1，x_2，…，x_n 来自指数分布 $\exp(\lambda)$，所以次序统计量 $x_{(1)}$，$x_{(2)}$，…，$x_{(n)}$ 的联合密度函数为：

$$p(t_1\cdots t_n)=n!\mathrm{e}^{-\lambda\sum_{i=1}^{n}t_i}$$

令 $z_1=nx_{(1)}$

$z_2=(n-1)(x_{(2)}-x_{(1)})$

$\vdots$

$z_n=x_{(n)}$

$$\therefore \quad |J|^{-1}=\begin{vmatrix} n & \cdots & & \\ & n-1 & & \\ & & \ddots & \\ & & & 1 \end{vmatrix}=n! \Rightarrow |J|=\frac{1}{n!}$$

$$\therefore \quad f(z_1,\cdots,z_n)=\mathrm{e}^{-\lambda\sum\limits_{i=1}^{n}z_i} \quad (\sum_{i=1}^{n}z_i=\sum_{i=1}^{n}t_i)$$

由该联合密度可知 z_1，…，z_n 独立同分布且 $z_i\sim\exp(\lambda)$。

13. $N(\ln2/\lambda，1/(n\lambda^2))$。

习题 1.5

1. （1）由因子分解定理可得。（2）略。（3）由定义可得。
2. 假设不失一般性，$a_1<a_2<\cdots<a_k$，则已知（n_1，n_2，…，n_k）后可得样本次序统计量为（$a_1\cdots a_1$，$a_2\cdots a_2$，…，$a_k\cdots a_k$）。
   因为次序统计量是充分统计量，
   所以可知（n_1，n_2，…，n_k）是充分统计量。
3. $T=\sum\limits_{i=1}^{n}x_i$，利用因子分解定理结论可证。
4. 因子分解定理 $T=(T_1，T_2)$，$T_1=\sum\limits_{i=1}^{n}\ln x_i$，$T_2=\sum\limits_{i=1}^{n}(\ln x_i)^2$。
5. $T=(T_1，T_2)$，$T_1=\sum\limits_{i=1}^{n}x_i$，$T_2=\prod\limits_{i=1}^{n}x_i$。
6. $T=(T_1，T_2)$，$T_1=\sum\limits_{i=1}^{n}\ln x_i$，$T_2=\sum\limits_{i=1}^{n}\ln(1-x_i)$。
7. （1）$T=\prod\limits_{i=1}^{n}x_i$。（2）$T=\prod\limits_{i=1}^{n}x_i$。（3）$T=\sum\limits_{i=1}^{n}|x_i|$。
8. （1）$T=x_{(n)}$。（2）$T=(x_{(1)}，x_{(n)})$。（3）$T=(x_{(1)}，x_{(n)})$。
9. 利用因子分解定理可证。
10. 利用因子分解定理可证。
11. $T=\sum\limits_{i=1}^{n}x_i^m$。
12. $T=x_{(1)}$。
13. $T=(\bar{x}，\bar{y}，\sum\limits_{i=1}^{n}(x_i-\bar{x})^2，\sum\limits_{i=1}^{n}(y_i-\bar{y})^2，\sum\limits_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}))$。

习题 1.6

1. 1/3，2/3，2/3。
2. 提示：先求 x_1/x_2 与 $\sqrt{x_1^2+x_2^2}$ 各自的分布，再求其联合分布，可证明联合分布是各自分布的乘积，所以独立。
3. $\mathrm{Cov}(\boldsymbol{y})=\mathrm{Cov}(\boldsymbol{Ax})=\boldsymbol{A}'\mathrm{Cov}(\boldsymbol{x})\boldsymbol{A}$

$\therefore \mathrm{Cov}(y)=\sigma^2 \mathrm{diag}(1+(n-1)\rho,1-\rho,\cdots,1-\rho)$

4.（1）求 Y_1 的分布。令 $X=X_1$，$Y=X_1+X_2$，则

$\begin{cases}X_1=X\\X_2=Y-X\end{cases}$ 且 $|J|=\begin{vmatrix}1 & -1\\0 & 1\end{vmatrix}=1$，则 x，y 的联合分布密度函数为：

$$p(x,y)=p(x_1,x_2)|J|=\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}x^{\alpha_1-1}\mathrm{e}^{-\lambda x}\frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}(y-x)^{\alpha_2-1}\mathrm{e}^{-\lambda(y-x)}$$

则 y 的边缘分布密度函数为：

$$\begin{aligned}p(y)&=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}x^{\alpha_1-1}\mathrm{e}^{-\lambda x}\ \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}(y-x)^{\alpha_2-1}\mathrm{e}^{-\lambda(y-x)}\mathrm{d}x\\&=\frac{\lambda^{\alpha_1+\alpha_2}\mathrm{e}^{-\lambda y}y^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1+\alpha_2)}\int_{-\infty}^{+\infty}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)+\Gamma(\alpha_2)}\left(\frac{x}{y}\right)^{\alpha_1-1}\left(1-\frac{x}{y}\right)^{\alpha_2-1}\mathrm{d}\,\frac{x}{y}\\&=\frac{\lambda^{\alpha_1+\alpha_2}\mathrm{e}^{-\lambda y}y^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1+\alpha_2)}\end{aligned}$$

（2）求 Y_2 的分布。令 $X=X_1$，$Y=X_1/(X_1+X_2)$，则

$\begin{cases}X_1=X\\X_2=\dfrac{X}{Y}-X\end{cases}$ 且 $|J|=\begin{vmatrix}1 & \dfrac{1}{y}-1\\0 & -\dfrac{x}{y^2}\end{vmatrix}=-\dfrac{x}{y^2}$，则 x，y 的联合分布密度函数为：

$$\begin{aligned}p(x,y)&=p(x_1,x_2)|J|\\&=\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}x^{\alpha_1-1}\mathrm{e}^{-\lambda x}\frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}\left(\frac{x}{y}-x\right)^{\alpha_2-1}\mathrm{e}^{-\lambda\left(\frac{x}{y}-x\right)}\left(-\frac{x}{y^2}\right)\end{aligned}$$

则 y 的边缘密度函数为：

$$\begin{aligned}p(y)&=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}x^{\alpha_1-1}\mathrm{e}^{-\lambda x}\frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}\left(\frac{x}{y}-x\right)^{\alpha_2-1}\cdot\mathrm{e}^{-\lambda\left(\frac{x}{y}-x\right)}\left(-\frac{x}{y^2}\right)\mathrm{d}x\\&=\frac{-\dfrac{1}{y^2}\left(\dfrac{1}{y}-1\right)^{\alpha_2-1}\Gamma(\alpha_1+\alpha_2)y^{\alpha_1+\alpha_2-1}y}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_{-\infty}^{+\infty}\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}\left(\frac{x}{y}\right)^{\alpha_1+\alpha_2-1}\mathrm{e}^{-\lambda\frac{x}{y}}\mathrm{d}\,\frac{x}{y}\\&=\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y^{\alpha_1-1}(1-y)^{\alpha_2-1}\end{aligned}$$

即 $Y_2\sim Be(\alpha_1，\alpha_2)$。

（3）Y_1 与 Y_2 独立证明：

Y_1 与 Y_2 的联合分布：令 $X=X_1+X_2$，$Y=X_1/(X_1+X_2)$

$\begin{cases}X_1=XY\\X_2=X-XY\end{cases}$ 且 $|J|=\begin{vmatrix}y & 1-y\\x & -x\end{vmatrix}=-x$

则 x，y 的联合分布密度函数为：

$$
\begin{aligned}
p(x,y)&=p(x_1,x_2)|J|\\
&=\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}(xy)^{\alpha_1-1}e^{-\lambda xy}\frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}(x-xy)^{\alpha_2-1}e^{-\lambda(x-xy)}(-x)\\
&=\frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}x^{\alpha_1+\alpha_2-1}e^{-\lambda x}\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y^{\alpha_1-1}(1-y)^{\alpha_2-1}\\
&=p(x)p(y)
\end{aligned}
$$

故独立。

5. （1）$X\sim Be(a,\ b)$，则 $p(X)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}\ (1-x)^{b-1}$，$0<x<1$。

$Y=\frac{X}{1-X}\Rightarrow X=\frac{Y}{1+Y}$单增

$$
\begin{aligned}
\therefore p(y)&=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{Y}{1+Y}\right)^{a-1}\left(1-\frac{Y}{1+Y}\right)^{b-1}\cdot(1+Y)^{-2}\\
&=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}Y^{a-1}(1+Y)^{-(a+b)}
\end{aligned}
$$

$\therefore Y\sim F(2a,\ 2b)$。

（2）见第6题证明。

6. $Z=\frac{n}{m}X\Big/\left(1+\frac{n}{m}X\right)\Rightarrow X=\frac{m}{n}\frac{z}{1-z}$单增

$$
\begin{aligned}
\therefore p(z)&=f_x\left(\frac{m}{n}\frac{z}{1-z}\right)\left|\left(\frac{m}{n}\frac{z}{1-z}\right)'\right|\\
&=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{n}{m}\right)^{\frac{n}{2}}\left(\frac{m}{n}\frac{z}{1-z}\right)^{\frac{n}{2}-1}\left(1+\frac{z}{1-z}\right)^{-\frac{m+n}{2}}\cdot\frac{m}{n}\cdot\frac{1}{(1-z)^2}\\
&=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(-1+\frac{1}{1-z}\right)^{\frac{n}{2}-1}\left(\frac{1}{1-z}\right)^{-\frac{m+n}{2}}\frac{1}{(1-z)^2}\\
&=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}z^{\frac{n}{2}-1}(1-z)^{\frac{m}{2}-1}
\end{aligned}
$$

故服从 $Be\left(\frac{n}{2},\ \frac{m}{2}\right)$。

（2）当 $X=\frac{bz}{a(1-z)}$时，可以使 X 服从 F 分布。

7. （1）是，$\sum_{i=1}^{n}x_i$。

（2）是，$\left(\sum_{i=1}^{n}\ln x_i,\sum_{i=1}^{n}(\ln x_i)^2\right)$。

（3）不是。

(4) 是，$\sum_{i=1}^{n} \mid x_i \mid$。

(5) 不是。

(6) 不是。

8. 是，$\left(\sum_{i=1}^{n} x_{1i}, \sum_{i=1}^{n} x_{2i}, \sum_{i=1}^{n} x_{1i}^2, \sum_{i=1}^{n} x_{2i}^2, \sum_{i=1}^{n} x_{1i}x_{2i}\right)$。

习题 2.1

1. ∵x_1，x_2，…，x_n 来自伽玛分布 $Ga(\alpha,\lambda)$，易知均值 $\mu=\frac{\alpha}{\lambda}$，方差 $\sigma^2=\frac{\alpha}{\lambda^2}$

∴$\alpha=\frac{\mu^2}{\sigma^2}$，$\lambda=\frac{\mu}{\sigma^2}$

由 $\hat{\mu}=\bar{x}$，$\hat{\sigma}^2=\hat{s}_n^2$ 知 α 和 λ 的矩估计为：

$\hat{\lambda}=\frac{\bar{x}}{s_n^2}$，$\hat{\alpha}=\frac{\bar{x}^2}{s_n^2}$

2. 因为 x_1，x_2，…，x_n 来自两点分布 $b(1,\theta)$，易知均值 $\mu=p$，方差 $\sigma^2=p(1-p)$。

所以 θ 与 $g(\theta)$ 和矩估计为：$\hat{\theta}=\bar{x}$，$\hat{g}_1(\theta)=\bar{x}(1-\bar{x})$。

3. 由题可知 $EX=\sum_{k=1}^{N}k\cdot\frac{1}{N}=\frac{N+1}{2}$

所以正整数 N 的矩估计为 $\hat{N}=2\bar{X}-1$。

4. 因为 x_1，x_2，…，x_n 是服从密度函数为 $P(x)=\sqrt{\theta}x^{\sqrt{\theta}-1}$ $0<x<1$ 的分布函数。所以

$$\mu=\int_0^1 x\cdot\sqrt{\theta}x^{\sqrt{\theta}-1}\mathrm{d}x=\sqrt{\theta}\int_0^1 x^{\sqrt{\theta}}\mathrm{d}x$$

$$=\sqrt{\theta}\frac{x^{\sqrt{\theta}+1}}{\sqrt{\theta}+1}\Big|_0^1=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$$

即 $\mu=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$ $(0<\mu<1)$，可得 $\theta=\frac{\mu^2}{(1-\mu)^2}$。

所以 θ 的矩估计为 $\hat{\theta}=\frac{\bar{x}^2}{(1-\bar{x})^2}$。

5. (1) 总字数为 N，令甲识别出错别字的概率为 p_1，乙识别出错别字的概率为 p_2，甲、乙独立，则同时识别出一个字的概率为 p_1p_2。

由矩估计的思想知：$\hat{p}_1=\frac{A}{N}$，$\hat{p}_2=\frac{B}{N}$，$\hat{p}_1\hat{p}_2=\frac{C}{N}$

甲、乙独立，可得到 $\frac{AB}{N^2}=\frac{C}{N}\Rightarrow\hat{N}=\frac{AB}{C}$。

(2) 未被甲、乙发现的错字数 M 的估计，为总错字数减去甲、乙发现的错字数，即 $\hat{M}=\hat{N}-A-B+C=\frac{AB}{C}-(A+B-C)$

当 $A=80$，$B=70$，$C=50$ 时，$\hat{N}=112$，$\hat{M}=12$。

6. 由 x_1，x_2，…，x_n 来自均匀分布 $U(0, \theta)$，易知 $\mu=\frac{\theta}{2}$，θ 的矩估计为 $\hat{\theta}=2\bar{x}$。因为

$E(\hat{\theta})=2E(\bar{x})=\theta$

所以具有无偏性。

证明相合性，即证 $P(|2\bar{x}-2\mu|>\varepsilon)\to 0\ (n\to\infty)$，即

$P\left(|\bar{x}-\mu|>\frac{\varepsilon}{2}\right)\to 0\ (n\to\infty)$

由辛钦大数定律知 θ 的矩估计满足相合性。

7. 证明：由切比雪夫不等式

$$P(|\hat{\theta}_n-\theta|>\varepsilon)\leqslant\frac{\mathrm{Var}(\hat{\theta}_n)}{\varepsilon^2}$$

由题知 $\lim\limits_{n\to\infty}E(\hat{\theta}_n)=\theta$，$\lim\limits_{n\to\infty}\mathrm{Var}(\hat{\theta}_n)=0$，所以当 $n\to\infty$时，$\frac{\mathrm{Var}(\hat{\theta}_n)}{\varepsilon^2}\to 0$，即 $P(|\hat{\theta}_n-\theta|>\varepsilon)\to 0\ (n\to\infty)$，即证。

8. 因为 x_1，x_2，…，x_n 来自泊松分布 $P(\lambda)$，可知 $EX=\lambda$。

$$P_0=P(X=0)=\mathrm{e}^{-\lambda}$$

所以 $P(x=0)$ 的矩估计为 $\hat{P}_0=\mathrm{e}^{-\bar{x}}$。

由辛钦大数定律易知 $\bar{x}$ 是λ 的相合估计，由定理 2.1.2

可知 $\mathrm{e}^{-\bar{x}}$是 $\mathrm{e}^{-\lambda}$的相合估计。

9. 因为 x_1，x_2，…，x_n 来自指数分布 $\exp(\lambda)$，易知 $EX=\frac{1}{\lambda}$。

所以 λ 的矩估计是 $\hat{\lambda}=\frac{1}{\bar{X}}$。

$$P(|\hat{\lambda}-\lambda|>\varepsilon)=P\left(\left|\frac{1}{\bar{X}}-\frac{1}{EX}\right|>\varepsilon\right)=P(|\bar{X}-EX|>\varepsilon\cdot\bar{X}\cdot EX)$$

ε 取任意小，由辛钦大数定律知具有相合性即 $P(|\hat{\lambda}-\lambda|>\varepsilon)\to 0\ (n\to\infty)$。

无偏性的判断：$T=\sum\limits_{i=1}^{n}x_i$ 服从 $Ga(n, \lambda)$（见 3.1.2 节例 3.1.3)，从而$\frac{1}{T}$服从倒伽玛分布 $IGa(n, \lambda)$，因为令 $Y=\frac{1}{T}$，$f_Y(y)=f_T(1/y)\left|\left(\frac{1}{y}\right)'\right|=\frac{\lambda^n}{\Gamma(n)}y^{-n-1}\mathrm{e}^{-\frac{\lambda}{y}}$

所以 $EY=\frac{\lambda}{n-1}$。

而 $\hat{\lambda}=\frac{1}{\bar{X}}=nY$，所以 $E\hat{\lambda}=\frac{\lambda n}{n-1}$。

故不是无偏估计。

习题 2.2

1. 总体 X 服从参数为λ 的分布，则（x_1，…，x_n）的联合密度函数为：

$$p(x_1,\cdots,x_n;\lambda)=\frac{\lambda^{\sum\limits_{i=1}^{n}k_i}}{x_1!\cdots x_n!}\mathrm{e}^{-n\lambda}$$

两边取对数得

$$\ln p(x_1,\cdots,x_n;\lambda)=\sum_{i=1}^{n}k_i\ln\lambda-n\lambda-\ln(x_1!\cdots x_n!)$$

对 λ 求导得

$$\frac{\partial\ln p(x_1,\cdots,x_n;\lambda)}{\partial\lambda}=\frac{\sum\limits_{i=1}^{n}k_i}{\lambda}-n$$

从而可得 λ 的最大似然估计为 $\hat{\lambda}=\frac{1}{n}\sum\limits_{i=1}^{n}k_i$。

2. (1) 总体 X 的密度函数为 $p(x;\ \beta)=(\beta+1)x^{\beta}$ ($0<x<1$)，则 (x_1，…，x_n) 的联合密度函数为：

$$p(x_1,\cdots,x_n;\beta)=(\beta+1)^n(x_1\cdots x_n)^{\beta}$$

两边取对数得

$$\ln p(x_1,\cdots,x_n;\beta)=n\ln(\beta+1)+\beta\ln(x_1\cdots x_n)$$

对 β 求导得

$$\frac{\partial\ln p(x_1,\cdots,x_n;\beta)}{\partial\beta}=\frac{n}{\beta+1}+\ln(x_1\cdots x_n)$$

从而 β 的最大似然估计 $\hat{\beta}=-\frac{n}{\ln(x_1\cdots x_n)}-1$ (因为 $\frac{\partial^2\ln p(x_1,\ \cdots,\ x_n;\ \beta)}{\partial\beta^2}<0$)。

矩估计 $\mu=EX=(\beta+1)\int_0^1 x\cdot x^{\beta+1}\mathrm{d}x=\frac{\beta+1}{\beta+2}$，所以 β 的矩估计 $\hat{\beta}=\frac{1}{1-\overline{X}}-2$，易知最大似然估计与矩法估计两者结果不一致。

(2) 给定样本观察值 $\bar{x}=0.515$，最大似然估计 $\hat{\beta}=0.39$，矩估计 $\hat{\beta}=0.061\,9$。

3. 总体 X 的密度函数为：

$$p(x,\ \sigma)=\frac{1}{2\sigma}\mathrm{e}^{-|x|/\sigma},\quad -\infty<x<+\infty$$

则 (x_1，…，x_n) 的联合密度函数为：

$$p(x_1,\cdots,x_n;\sigma)=\frac{1}{(2\sigma)^n}\mathrm{e}^{-\frac{\sum\limits_{i=1}^{n}|x_i|}{\sigma}}$$

两边取对数得

$$\ln p(x_1,\cdots,x_n;\sigma)=-\frac{\sum_{i=1}^{n}|x_i|}{\sigma}-n\ln(2\sigma)$$

对 σ 求导：

$$\frac{\partial \ln p(x_1,\cdots,x_n;\sigma)}{\partial\sigma}=\frac{\sum_{i=1}^{n}|x_i|}{\sigma^2}-\frac{n}{\sigma}$$

从而可得 σ 的最大似然估计 $\hat{\sigma}=\frac{1}{n}\sum_{i=1}^{n}|x_i|$。

4. 由题可知

$$p(x_1,\cdots,x_n;y_1,\cdots,y_m;\mu_1,\mu_2,\sigma^2)$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{m+n}{2}}\exp\left[-\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2+\sum_{i=1}^{m}(y_i-\mu_2)^2}{2\sigma^2}\right]$$

两边取对数得

$$\ln p(X,Y,\mu_1,\mu_2,\sigma^2)=-\frac{m+n}{2}\ln(2\pi\sigma^2)-\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2+\sum_{i=1}^{m}(y_i-\mu_2)^2}{2\sigma^2}$$

令 $\frac{\partial \ln l}{\partial \mu_1}=0\Rightarrow\hat{\mu}_1=\bar{x}$

令 $\frac{\partial \ln l}{\partial \mu_2}=0\Rightarrow\hat{\mu}_2=\bar{y}$

令 $\frac{\partial \ln l}{\partial \sigma^2}=0\Rightarrow\hat{\sigma}^2=\frac{1}{n+m}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{i=1}^{m}(y_i-\bar{y})^2\right]$

5. 当 $m=2$ 时，该截尾分布只能取1与2，设 x_1，…，x_n 的样本中有 n_1 个 x_i 为1，有 $n-n_1$ 个 x_i 为2，则似然函数为：

$$L(p)=\frac{p^{n_1}(1-p)^{n_1}p^{2(n-n_1)}}{[1-(1-p)^2]^n}=\frac{p^{2n-n_1}(1-p)^{n_1}}{[1-(1-p)^2]^n}=\frac{p^{n-n_1}(1-p)^{n_1}}{(2-p)^n}$$

$$\ln L(\theta)=(n-n_1)\ln p+n_1\ln(1-p)-n\ln(2-p)$$

令 $\frac{\partial \ln L(\theta)}{\partial p}=0\Rightarrow\hat{p}=\frac{2(n-n_1)}{2n-n_1}=\frac{2(\bar{x}-1)}{\bar{x}}$

$\left(\sum_{i=1}^{n}x_i=n_1+2(n-n_1)=n\bar{x}\right)$

6. (1) 总体 X 服从几何分布 $P(X=k)=p(1-p)^{k-1}$，$k=1$，2，…

$$p(x_1,\cdots,x_n;p)=p^n(1-p)^{\sum_{i=1}^{n}x_i-n}$$

则 $\ln p(x_1,\cdots,x_n;p)=n\ln p+\left(\sum_{i=1}^{n}x_i-n\right)\ln(1-p)$

$$\frac{\partial \ln p(x_1,\cdots,x_n;p)}{\partial p}=\frac{n}{p}-\frac{\sum_{i=1}^{n}x_i-n}{1-p}\quad 且\frac{\partial^2 \ln p(x_1\cdots x_n,\ p)}{\partial p^2}<0$$

$\Rightarrow p$ 的最大似然估计 $\hat{p}=\dfrac{n}{\sum_{i=1}^{n}x_i}=\dfrac{1}{\bar{x}}$

(2) $EX=\sum_{k=1}^{\infty}kp(1-p)^{k-1}=\dfrac{1}{p}$

由最大似然估计的不变性知 EX 的最大似然估计为 $\bar{x}$。

7. (1) 由 2.2.2 节例 2.2.7 可知，μ 的 MLE 为 $\hat{\mu}=\bar{x}$，σ 的 MLE 为 $\hat{\sigma}=\sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$。

由题知 $X\sim N(\mu,\ \sigma^2)$，可得 $\dfrac{X-\mu}{\sigma}\sim N(0,\ 1)$。

$$P(X>A)=P\left(\frac{X-\mu}{\sigma}>\frac{A-\mu}{\sigma}\right)=0.05$$

查表可得 $\dfrac{A-\mu}{\sigma}=1.645\Rightarrow A=\mu+1.645\sigma$

由最大似然估计的不变性知 $\hat{A}=\hat{\mu}+1.645\hat{\sigma}$。

(2) $\theta=P(x\geqslant 2)=P\left(\dfrac{X-\mu}{\sigma}\geqslant\dfrac{2-\mu}{\sigma}\right)=1-\Phi\left(\dfrac{2-\mu}{\sigma}\right)$

由最大似然估计的不变性知 $\hat{\theta}=1-\Phi\left(\dfrac{2-\hat{\mu}}{\hat{\sigma}}\right)$。

8. 样本来自于 Pareto 分布，故

$$p(x_1,\cdots,x_n;\alpha,\theta)=\theta^n\alpha^{n\theta}\left(\prod_{i=1}^{n}x_i\right)^{-(\theta+1)},\quad x>\alpha>0,\theta>0$$

两边求对数得

$$\ln p(x_1,\cdots,x_n;\alpha,\theta)=n\ln\theta+n\theta\ln\alpha-(\theta+1)\ln\prod_{i=1}^{n}x_i$$

分别对 α 与 θ 求导：

$$\frac{\partial \ln p(x_1,\cdots,x_n;\alpha,\theta)}{\partial\alpha}=\frac{n\theta}{\alpha}>0$$

$$\frac{\partial \ln p(x_1,\cdots,x_n;\alpha,\theta)}{\partial\theta}=\frac{n}{\theta}+n\ln\alpha-\ln\prod_{i=1}^{n}x_i=0$$

由 $\dfrac{\partial \ln p(x_1,\cdots,x_n;\alpha,\theta)}{\partial\alpha}>0$ 而 $x>\alpha>0$ 可知，α 的 MLE 为 $\hat{\alpha}=x_{(1)}$。

由 $\dfrac{\partial \ln p(x_1,\ \cdots,\ x_n;\ \alpha,\ \theta)}{\partial\theta}=0$ 得 $\theta=\dfrac{n}{\sum_{i=1}^{n}\ln x_i-n\ln\alpha}$ 且 $\dfrac{\partial^2 \ln p(x_1,\ \cdots,\ x_n;\ \theta,\ \alpha)}{\partial\theta^2}<0$，

可得 θ 的 MLE 为 $\hat{\theta}=\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i-n\ln x_{(1)}}$。

9. 证明：$p(x;\theta)$ 为 Cramer-Rao 正则族分布，且其二阶偏导数存在。

$\because 1=\int_{-\infty}^{+\infty}p(x;\theta)\mathrm{d}x$

对 θ 求导得

$$0=\int_{-\infty}^{+\infty}\frac{\partial\, p(x;\theta)}{\partial\,\theta}\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{\partial\,\ln p(x;\theta)}{\partial\,\theta}p(x;\theta)\mathrm{d}x$$

即 $E\dfrac{\partial\,\ln p(x;\theta)}{\partial\,\theta}=0$。

两对 θ 求导得

$$\begin{aligned}0&=\int_{-\infty}^{+\infty}\left[\frac{\partial^2\,\ln p(x,\theta)}{\partial\theta^2}p(x;\theta)+\frac{\partial\ln p(x;\theta)}{\partial\theta}\,\frac{\partial p(x;\theta)}{\partial\theta}\right]\mathrm{d}x\\&=\int_{-\infty}^{+\infty}\left[\frac{\partial^2\,\ln p(x;\theta)}{\partial\theta^2}p(x;\theta)+\left(\frac{\partial\ln p(x;\theta)}{\partial\theta}\right)^2p(x;\theta)\right]\mathrm{d}x\end{aligned}$$

从而
$$\begin{aligned}I(\theta)&=E\left(\frac{\partial\,\ln p(x;\theta)}{\partial\,\theta}\right)^2=\int_{-\infty}^{+\infty}\left(\frac{\partial\,\ln p(x;\theta)}{\partial\,\theta}\right)^2p(x;\theta)\mathrm{d}x\\&=-\int_{-\infty}^{+\infty}\frac{\partial^{\,2}\,\ln p(x;\theta)}{\partial\,\theta^2}p(x;\theta)\mathrm{d}x\\&=-E\left(\frac{\partial^{\,2}\,\ln p(x;\theta)}{\partial\,\theta^2}\right)\end{aligned}$$

即证。

10. (1) z 总体服从瑞利分布：

$$p(z_1,\cdots,z_n;\sigma^2)=\frac{\prod\limits_{i=1}^{n}z_i}{\sigma^{2n}}\exp\left(-\frac{\sum\limits_{i=1}^{n}z_i^2}{2\sigma^2}\right)$$

$$\frac{\partial\ln p(z_1,\cdots,z_n;\sigma^2)}{\partial\sigma^2}=0$$

可得 $\hat{\sigma}^2_{\mathrm{MLE}}=\dfrac{1}{2n}\sum\limits_{i=1}^{n}z_i^2$，$E(\hat{\sigma}^2_{\mathrm{MLE}})=\dfrac{1}{2n}\sum\limits_{i=1}^{n}E(z_i^2)=\sigma^2$，故是 σ^2 的无偏估计。

(2) $\ln p(z)=\ln z-\dfrac{z^2}{2\sigma^2}-\ln\sigma^2$

$$\frac{\partial\ln p(z)}{\partial\sigma^2}=\frac{z^2}{2\sigma^4}-\frac{1}{\sigma^2}=\frac{z^2-2\sigma^2}{2\sigma^4}$$

$$\frac{\partial^2\,\ln p(z)}{\partial(\sigma^2)^2}=\frac{1}{\sigma^4}-\frac{z^2}{\sigma^6}$$

$$\begin{aligned}\therefore\ I(\sigma^2)&=-E\left(\frac{\partial^{\,2}\,\ln p(z)}{\partial\,(\sigma^2)^2}\right)=-E\left(\frac{1}{\sigma^4}-\frac{z^2}{\sigma^6}\right)\\&=-\frac{1}{\sigma^4}+E\left(\frac{x^2+y^2}{\sigma^6}\right)\end{aligned}$$

$$=-\frac{1}{\sigma^4}+\frac{2\sigma^2}{\sigma^6}=\frac{1}{\sigma^4}$$

（3）由定理 2.2.2 知

$$\hat{\sigma}^2 \sim AN(\sigma^2, \frac{\sigma^4}{n})$$

11. 步骤同第 10 题：（1）$\hat{\sigma}^2=\frac{1}{2n}\sum_{i=1}^{n} z_i^{-2}$；（2）$I(\sigma^2)=\frac{1}{\sigma^4}$；（3）$\hat{\sigma}^2 \sim AN\left(\sigma^2, \frac{\sigma^4}{n}\right)$。

12. （1）总体服从伽玛分布 $Ga(\alpha, \lambda)$，α 已知，则

$$p(x_1,\cdots,x_n,\lambda)=\frac{\lambda^{n\alpha}}{\Gamma^n(\alpha)}\left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} e^{-\lambda\sum_{i=1}^{n} x_i}$$

取对数得

$$\ln p(x_1,\cdots,x_n,\lambda)=n\alpha\ln\lambda+(\alpha-1)\ln\prod_{i=1}^{n} x_i-\lambda\sum_{i=1}^{n} x_i-n\ln\Gamma(\alpha)$$

由$\frac{\partial\ln p(x_1,\cdots,x_n,\lambda)}{\partial\lambda}=0$ 知 λ 的最大似然估计 $\hat{\lambda}=\frac{\alpha}{\bar{x}}$。

（2）求其渐近分布，首先求出费歇信息量：

$$\ln p(x;\lambda)=\alpha\ln\lambda+(\alpha-1)\ln x-\lambda x-\ln\Gamma(\alpha)$$

$$\frac{\partial\ln p(x;\lambda)}{\partial\lambda}=\frac{\alpha}{\lambda}-x$$

$$\frac{\partial^2\ln p(x;\lambda)}{\partial\lambda^2}=-\frac{\alpha}{\lambda^2}$$

$$\Rightarrow I(\lambda)=-E\left(\frac{\partial^2\ln p(x;\lambda)}{\partial\lambda^2}\right)=E\left(\frac{\alpha}{\lambda^2}\right)=\frac{\alpha}{\lambda^2}$$

所以由定理 2.2.2 知，$\hat{\lambda}\sim AN(\lambda, \frac{\lambda^2}{n\alpha})$。

13. $\hat{\theta}_M=\frac{9}{32}$，$\hat{\theta}=\frac{8-4\sqrt{3}}{13}$。

14. 188 小时。

习题 2.3

1. $\because Ex_1=Ex_2=Ex_3=\mu \quad \therefore E\hat{\mu}_1=E\hat{\mu}_2=E\hat{\mu}_3=\mu$，即都是 μ 的无偏估计。

$$\because \mathrm{Var}(\hat{\mu}_1)=\frac{1}{4}\mathrm{Var}(x_1)+\frac{1}{9}\mathrm{Var}(x_2)+\frac{1}{36}\mathrm{Var}(x_3)=\frac{14}{36}\sigma^2$$

$$\mathrm{Var}(\hat{\mu}_2)=\frac{1}{9}\mathrm{Var}(x_1)+\frac{1}{9}\mathrm{Var}(x_2)+\frac{1}{9}\mathrm{Var}(x_3)=\frac{12}{36}\sigma^2$$

$$\mathrm{Var}(\hat{\mu}_3)=\frac{1}{36}\mathrm{Var}(x_1)+\frac{1}{36}\mathrm{Var}(x_2)+\frac{4}{9}\mathrm{Var}(x_3)=\frac{18}{36}\sigma^2$$

故 $\hat{\mu}_2$ 最有效。

2. （1）$E(\hat{\theta}_\alpha)=E(\alpha\hat{\theta}_1)+E[(1-\alpha)\hat{\theta}_2]$

$$=\alpha\theta+(1-\alpha)\theta=\theta$$

故为无偏估计。

(2) $\hat{\theta}_1$ 与 $\hat{\theta}_2$ 相互独立。

$$\begin{aligned}\mathrm{Var}(\hat{\theta}_\alpha)&=\mathrm{Var}(\alpha\hat{\theta}_1)+\mathrm{Var}[(1-\alpha)\hat{\theta}_2]\\&=\alpha^2\sigma_1^2+(1-\alpha)^2\sigma_2^2\\&=(\sigma_1^2+\sigma_2^2)\alpha^2-2\alpha\sigma_2^2+\sigma_2^2\end{aligned}$$

易知当 $\alpha=\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ 时，$\mathrm{Var}(\hat{\theta}_\alpha)$ 取最小值。

3. (1) 由题可知 $\bar{X}\sim N\left(\mu_1,\ \frac{1}{n}\right)$，$\bar{Y}\sim N\left(\mu_2,\ \frac{4}{m}\right)$，两样本独立，所以

$$\bar{X}-\bar{Y}\sim N\left(\mu_1-\mu_2,\ \frac{1}{n}+\frac{4}{m}\right)$$

$\mu=\mu_1-\mu_2$ 的无偏估计 $\hat{\mu}=\bar{X}-\bar{Y}$。

(2) $\mathrm{Var}(\hat{\mu})=\mathrm{Var}(\bar{X}-\bar{Y})$

$$=\frac{1}{n}+\frac{4}{m}\geqslant 2\sqrt{\frac{4}{n\cdot m}}$$

所以最小方差在 $\frac{1}{n}=\frac{4}{m}$ 即 $m=4n$ 时达到，为 $\frac{2}{\sqrt{n}}$。

4. 由题意可知 $Ex_i=\theta$，$E\hat{\theta}=\sum_{i=1}^{k}c_iEx_i=\theta\sum_{i=1}^{k}c_i$。

所以当 $\sum_{i=1}^{k}c_i=1$ 时，$\hat{\theta}$ 是 θ 的无偏估计。

k 台仪器各自独立地测量某物理量 θ 各一次，可知 x_i 之间独立。

$$\mathrm{Var}(\hat{\theta})=\sum_{i=1}^{k}c_i^2\mathrm{Var}(x_i)=\sum_{i=1}^{k}c_i^2\sigma_i^2$$

要使 $\mathrm{Var}(\hat{\theta})$ 达到最小，即对 $\sum_{i=1}^{k}c_i^2\sigma_i^2$ 在条件 $\sum_{i=1}^{k}c_i=1$ 下求条件极值，应用拉格朗日乘子法可得

$$C_i=\sigma_i^{-2}\Big/\left(\sum_{i=1}^{k}\sigma_i^{-2}\right),\ i=1,\ \cdots,\ k$$

5. (1) $E\hat{\theta}_1=E\bar{x}-\frac{1}{2}=\frac{\theta+\theta+1}{2}-\frac{1}{2}=\theta$

$$\because Ex_{(1)}=\int_\theta^{\theta+1}\{-xn[1-(x-\theta)]^{n-1}\}\mathrm{d}x$$

$$=\theta+\frac{1}{n+1}$$

$$\therefore E\hat{\theta}_2=Ex_{(1)}-\frac{1}{n+1}=\theta$$

$\because Ex_{(n)}=\int_{\theta}^{\theta+1}xn(x-\theta)^{n-1}\mathrm{d}x$

$=\theta+\dfrac{n}{n+1}$

$\therefore E\hat{\theta}_3=Ex_{(n)}-\dfrac{n}{n+1}=\theta$

故 $\hat{\theta}_1$，$\hat{\theta}_2$ 与 $\hat{\theta}_3$ 都是 θ 的无偏估计。

(2) $\mathrm{Var}(\hat{\theta}_1)=\mathrm{Var}(\bar{x})=\dfrac{1}{12n}$，$\mathrm{Var}(x_{(1)})=\mathrm{Var}(x_{(n)})=\dfrac{2}{(n+1)(n+2)}-\dfrac{1}{(n+1)^2}$

故当 $1\leqslant n\leqslant 7$ 时，$\hat{\theta}_1$ 比 $\hat{\theta}_2$，$\hat{\theta}_3$ 有效；当 $n\geqslant 8$ 时，$\hat{\theta}_2$，$\hat{\theta}_3$ 比 $\hat{\theta}_1$ 有效。

6. (1) 同第 5 题。

(2) 当 $n=1$，2 时，两个估计方差相等；当 $n>2$ 时，$\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效。

7. 总体来自指数分布 $\exp\left(\dfrac{1}{\theta}\right)$，所以，

$E\hat{\theta}_1=E\bar{x}=1\Big/\dfrac{1}{\theta}=\theta$

$E\hat{\theta}_2=nEx_{(1)}=n\int_0^{\infty}x(-n\lambda)\mathrm{e}^{-\lambda xn}\mathrm{d}x=\theta$

$\because \mathrm{Var}(\hat{\theta}_1)=\dfrac{\theta^2}{n}$

$\mathrm{Var}(\hat{\theta}_2)=\theta^2$

$\therefore \bar{x}$ 比 $\hat{\theta}_2$ 更有效。

8. $\because \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2}\sim\chi^2(n-1)$

$\therefore E\sum_{i=1}^{n}(x_i-\bar{x})^2=(n-1)\sigma^2$

$\mathrm{Var}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]=2(n-1)\sigma^4$

$$\begin{aligned}MSE(cQ)&=c^2\mathrm{Var}Q+(cEQ-\sigma^2)^2\\&=2(n-1)c^2\sigma^4+[c(n-1)\sigma^2-\sigma^2]^2\\&=2(n-1)c^2\sigma^4+[c(n-1)-1]^2\sigma^4\end{aligned}$$

解得当 $c=\dfrac{1}{n+1}$时，$MSE(cQ)$ 最小。

9. $MSE=E(\hat{\theta}-\theta)^2=\mathrm{Var}(\hat{\theta})+(E\hat{\theta}-\theta)^2$

故本题中 $MSE=\mathrm{Var}(c\bar{x})+(cE\bar{x}-\theta)^2$

$$\begin{aligned}&=c^2\theta^2/n+(c\theta-\theta)^2\\&=\left(\frac{1}{n}+1\right)c^2\theta^2-2c\theta^2+\theta^2\end{aligned}$$

$\therefore$ 当 $c=\dfrac{n}{n+1}$时，$c\bar{x}$ 在均方误差准则下为 θ 的最优估计。

10. 证明：由重期望公式可得

$E[E(\varphi(x)|T)]=E(\varphi|x)=g(\theta)$

故 $E(\varphi(x)|T)$ 可以作为 $g(\theta)$ 的一个估计。

因为 $T(x)$ 是充分统计量，故 $E(\varphi(x)\mid T)$ 与 θ 无关，且 $E(\varphi(x)\mid T)$ 是统计量。

$$\begin{aligned}\mathrm{Var}[\varphi(x)]&=E[\varphi(x)-g(\theta)]^2\\&=E\{\varphi(x)-E[\varphi(x)|T]+E[\varphi(x)|T]-g(\theta)\}^2\\&=E\{\varphi(x)-E[\varphi(x)|T]\}^2+E\{E[\varphi(x)|T]-g(\theta)\}^2\\&\geqslant\mathrm{Var}\{E[\varphi(x)|T]\}\end{aligned}$$

其中交叉项 $E\{\{\varphi(x)-E[\varphi(x)|T]\}\{E[\varphi(x)|T]-g(\theta)\}\}$

$$\begin{aligned}&=E\{E\{\{\varphi(x)-E[\varphi(x)|T]\}\{E[\varphi(x)|T]-g(\theta)\}|T(x)\}\}\\&=E\{\{E[\varphi(x)|T]-g(\theta)\}E\{\varphi(x)-E[\varphi(x)|T]|T(x)\}\}\\&=0\end{aligned}$$

而$MSE\{E[\varphi(x)|T]\}=\mathrm{Var}\{E[\varphi(x)|T]\}+\{E\{E[\varphi(x)|T]\}-g(\theta)\}^2$

$$=\mathrm{Var}\{E[\varphi(x)|T]\}$$

$$\begin{aligned}MSE[\varphi(x)]&=\mathrm{Var}[\varphi(x)]+[E\varphi(x)-g(\theta)]^2\\&=\mathrm{Var}[\varphi(x)]\end{aligned}$$

故 $E[\varphi(x)\mid T]$ 的均方误差不会超过 $\varphi(x)$ 的均方误差。

11. 反证：假设存在统计量 $\hat{g}(x)$ 是$\frac{1}{\theta}$的无偏估计，则

$$\int_0^\theta \hat{g}(x)\frac{1}{\theta}\mathrm{d}x=\frac{1}{\theta}$$

$$\therefore\int_0^\theta \hat{g}(x)\mathrm{d}x=1$$

$\therefore \hat{g}(x)=\frac{1}{\theta}$，不是统计量。

$\therefore\frac{1}{\theta}$的无偏估计不存在。

12. (1) $P_\lambda(k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$，$k=0$，1，…

$$E_\lambda\varphi(k)=\sum_{k=0}^{\infty}\varphi(k)\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$$

当 $E_\lambda\varphi(k)=0$ 时，$\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}>0$

因为只有当 $\varphi(k)=0$ 时 $E_\lambda\varphi(k)=0$，所以完备。

(2) $p_k=(1-p)^{k-1}p$，$k=1$，2，…

$$E\varphi(k)=\sum_{k=1}^{+\infty}\varphi(k)(1-p)^{k-1}p$$

$E\varphi(k)=0$ 时，$(1-p)^{k-1}p>0$

因为只有当 $\varphi(k)=0$ 时有 $E_x\varphi(k)=0$，所以是完备的。

（3）$p_t(x)=\dfrac{1}{\theta}$，$\theta>0$

$$E\varphi(x)=\int_0^\theta \varphi(x)\cdot\frac{1}{\theta}\mathrm{d}x$$

当 $E\varphi(x)=0$ 时，$\dfrac{1}{\theta}>0$。

易知 $\varphi(x)=0$，故完备。

（4）$p(x)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x}$，$x\geqslant 0$

$$E\varphi(x)=\int_0^{+\infty}\varphi(x)\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x}\mathrm{d}x$$

$$E\varphi(x)=0 \quad 即 \quad \int_0^{+\infty}\varphi(x)\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\lambda x}\mathrm{d}x=0$$

上式左边是函数 $\varphi(x)\cdot x^{\alpha-1}$ 的单边拉普拉斯变换，由单边拉普拉斯的唯一性可知 $P\{\varphi(x)x^{\alpha-1}=0\}=1$。$x>0$，故 $x^{\alpha-1}>0$，故只有 $\varphi(x)=0$，所以是完备统计量。

13. 对于分布为 $N(\mu,1)$ 的正态总体，$\sum\limits_{i=1}^n x_i$ 为其充分统计量，又因为该分布属于指数型分布族的分布，所以 $T_n=\sum\limits_{i=1}^n x_i$ 为完备充分统计量。

令 $\varphi(x_1,\cdots,x_n)=\begin{cases}1, & x_1\leqslant a\\ 0, & x_1>a\end{cases}$，容易看出 $\varphi(x)$ 是 $p=P(x_1\leqslant a)$ 的无偏估计。

所以 p 的 UMVUE 为：

$$E[\varphi(x_1,\cdots,x_n)\mid T_n]=P(x_1\leqslant a\mid T_n)$$
$$=P\left\{\frac{\sqrt{n}(x_1-\bar{x})}{\sqrt{n-1}\cdot s}\leqslant\frac{\sqrt{n}(a-\bar{x})}{\sqrt{n-1}\cdot s}\,\middle|\,\bar{x}\right\}$$

式中，$s=1$。

记 $u=\dfrac{\sqrt{n}\ (x_1-\bar{x})}{\sqrt{n-1}\cdot s}$，$\dfrac{Y_1}{\sqrt{n}}=\bar{x}$，$Y_2=\dfrac{\sqrt{n}(x_1-\bar{x})}{\sqrt{n-1}\cdot s}$，则

$$E[\varphi(x_1,\cdots,x_n)\mid T_n]=P(u\leqslant u_0\mid \bar{x})$$

当给定 $\bar{x}$ 和 S 时，u_0 为常数。

作正交变换 $\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}$，$\boldsymbol{A}=\begin{pmatrix}\dfrac{1}{\sqrt{n}} & \dfrac{1}{\sqrt{n}} & \cdots & \dfrac{1}{\sqrt{n}}\\ \dfrac{n-1}{\sqrt{n}\sqrt{n-1}} & \dfrac{-1}{\sqrt{n}\sqrt{n-1}} & \cdots & \dfrac{-1}{\sqrt{n}\sqrt{n-1}}\\ & c_{ij} & & \end{pmatrix}$

$\therefore \mathrm{Var}(\boldsymbol{Y})=\boldsymbol{A}\mathrm{Cov}(X)\boldsymbol{A}'=\sigma^2\boldsymbol{I}$

$\therefore Y_1$ 与 Y_2 独立

又∵ $\bar{x}_1-\bar{x}=\frac{n-1}{n}x_1-\frac{1}{n}\sum_{i=1}^{n}x_i\sim N(0,\frac{n-1}{n})$

$\therefore \frac{\sqrt{n}\ (x_1-\bar{x})}{\sqrt{n-1}}\sim N(0,\ 1)$

$\therefore E[\varphi(x_1,\cdots,x_n)\,|\,T_n]=P\left\{\frac{\sqrt{n}(x_1-\bar{x})}{\sqrt{n-1}\cdot S}\leqslant\frac{\sqrt{n}(a-\bar{x})}{\sqrt{n-1}}\right\}$

$$=\Phi\left[\sqrt{\frac{n}{n-1}}(a-\bar{x})\right]$$

$\therefore P$ 的 UMVUE 为 $\Phi\left[\sqrt{\frac{n}{n-1}}(a-\bar{x})\right]$

14\. ∵$X\sim\exp(\lambda)$，即 $X\sim Ga(1,\ \lambda)$

$\therefore\ T_n=\sum_{i=1}^{n}x_i\sim Ga(n,\lambda)$ 为完备充分统计量

取 $\varphi(x_1,\ \cdots,\ x_n)=\begin{cases}1, & x_1\leqslant a\\ 0, & x_1>a\end{cases}$

则 $\varphi(x_1,\ \cdots,\ x_n)$ 是 $p=P(x_1\leqslant a)$ 的无偏估计。

$\therefore E(\varphi(x_1\cdots x_n)\ |\ T_n)$ 为 p 的 UMVUE：

$E(\varphi(x_1\cdots x_n)\ |\ T_n)=P(x_1\leqslant a\ |\ T_n)$

若 $T_n=\sum_{i=1}^{n}x_i<a$，则必有 $x_1<a$，则 $E(\varphi(x_1\cdots x_n)\ |\ T_n)=1$

当 $T_n=\sum_{i=1}^{n}x_i\geqslant a$ 时，有

$E[\varphi(x_1\cdots x_n)\ |\ T_n]=P(x_1\leqslant a\ |\ T_n)=P\left(\frac{x_1}{T}\leqslant\frac{a}{T}\ |\ T_n\right)$

且 $\frac{x_1}{T}=\frac{\lambda x_1}{\lambda T}\sim\frac{Ga(1,\ 1)}{Ga(n,\ 1)}$，其分布与 λ 无关。

$\therefore\frac{x_1}{T}$ 为辅助统计量

又 T 为充分统计量

由 Basu 定理，可知 $\frac{x_1}{T}$ 与 T 相互独立。

$\therefore\frac{x_1}{T}=\frac{x_1}{x_1+x'_2}$，其中 $x'_2=x_2+\cdots+x_n\sim Ga(n-1,\ \lambda)$，$x_1\sim Ga(1,\ \lambda)$

$\therefore\frac{x_1}{T}\sim Be(1,\ n-1)$

$\therefore E[\varphi(x_1\cdots x_n)\,|\,T_n]=1-\left\{[1-\frac{a}{T}]^t\right\}^{n-1}$

其中 $T=\sum_{i=1}^{n}x_i$，$b^t=\begin{cases}b, & b\geqslant 0\\ 0, & b<0\end{cases}$

补充：定义：设 $X\sim f(x,\ \theta)$，$\theta\in\Theta$，若统计量 $A=A(X)$ 的分布与 θ 无关，则称

$A(X)$ 为辅助统计量，即 $A(X)$ 中不包含与 θ 有关的信息。

Basu 定理：设 $x\sim\{f(x,\theta),\theta\in\Theta\}$，$T=T(X)$ 为完备充分统计量，$A=A(X)$ 为辅助量，则 $T(X)$ 与 $A(X)$ 独立。证明略。

15. (1) 由于 $T=\sum_{i=1}^{n}x_i$ 是λ 的充分统计量，由伽玛分布的性质知 T 的密度函数

$$f(t)=\begin{cases}\dfrac{1}{\Gamma(n\alpha)}\lambda^{n\alpha}t^{n\alpha-1}\mathrm{e}^{-\lambda t}, & t>0\\ 0, & t\leqslant 0\end{cases}$$

$$E\left(\frac{1}{T}\right)=\int_0^{\infty}\frac{1}{\Gamma(n\alpha)}\lambda^{n\alpha}t^{n\alpha-2}\mathrm{e}^{-\lambda t}\,\mathrm{d}t=\frac{\lambda}{n\alpha-1}$$

令 $d(T)=\dfrac{n\alpha-1}{T}$，则 $E[d(T)]=\lambda$，从而

$$0=\int_0^{\infty}[\varphi_1(t)-\varphi_2(t)]\frac{1}{\Gamma(n\alpha)}\lambda^{n\alpha}t^{n\alpha-1}\mathrm{e}^{-\lambda t}\,\mathrm{d}t$$

即 $\int_0^{\infty}[\varphi_1(t)-\varphi_2(t)]\lambda^{n\alpha}t^{n\alpha-1}\mathrm{e}^{-\lambda t}\,\mathrm{d}t=0$

这是 $[\varphi_1(t)-\varphi_2(t)]t^{n\alpha-1}$的 Laplace 变换，由其象函数为 0 知 $[\varphi_1(t)-\varphi_2(t)]t^{n\alpha-1}=0$，对 $t>0$，从而 $\varphi_1(t)=\varphi_2(t)$，故 $d(T)=\dfrac{n\alpha-1}{T}$是$\lambda$ 的 UMVUE。

(2) $\because ET=\dfrac{n\alpha}{\lambda}$

令 $b(T)=\dfrac{T}{n\alpha}$，则有

$$E[b(T)]=\frac{1}{\lambda}$$

$$\mathrm{Var}[b(T)]=\frac{1}{\alpha^2}\cdot\frac{1}{n}\cdot\frac{\alpha}{\lambda^2}=\frac{1}{n\alpha\lambda^2}$$

又由于 $I(\lambda)=\dfrac{\alpha}{\lambda^2}$，$g(\lambda)=\dfrac{1}{\lambda}$，可得$\dfrac{1}{\lambda}$的 C-R 下界为$\dfrac{1}{n\alpha\lambda^2}$，故 λ^{-1}的 UMVUE 为$\dfrac{T}{n\alpha}$。

16. (1) 由于 T_1，T_2 分别是 θ_1，θ_2 的 UMVUE，故 $ET_i=\theta_i$，且对任意一个 $\phi(x)$ 满足 $E\phi=0$，由于 $\mathrm{Cov}(T_i,\phi)=0$ $(i=1,2)$。于是

$$E(aT_1+bT_2)=a\theta_1+b\theta_2$$

$$\mathrm{Cov}(aT_1+bT_2,\phi)=a\mathrm{Cov}(T_1,\phi)+b\mathrm{Cov}(T_2,\phi)=0$$

因此 aT_1+bT_2 是 $a\theta_1+b\theta_2$ 的 UMVUE。

(2) 正态总体下，$3\mu+4\sigma^2$ 的 UMVUE 为 $3\bar{x}+4s^2$（$\bar{x}$ 为样本均值，s^2 为样本方差）。

17. 因为 T 是$g(\theta)$ 的 UMVUE，$\hat{g}$ 是 $g(\theta)$ 的无偏估计，故其差 $T-\hat{g}$ 是 0 的无偏估计，即 $E(T-\hat{g})=0$，且 $D(T-\hat{g})<+\infty$，由于 $\mathrm{Cov}(T,T-\hat{g})=0$，所以

$D(T)-\mathrm{Cov}(T,\ \hat{g})=0$，即 $\mathrm{Cov}(T,\ \hat{g})=D(T)\geqslant 0$。

18. 略。

习题 2.4

1. (1) θ 的最大似然估计 $\hat{\theta}=-\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i}$，由似然估计的不变性知 $g(\theta)=1/\theta$ 的最大似然估计为：

$$\hat{g}(x)=-\frac{1}{n}\sum_{i=1}^{n}\ln x_i$$

(2) x 的密度函数为 $p(x;\ \theta)=\theta x^{\theta-1}$，则 $\ln x$ 服从指数分布 $\exp(\theta)$。

$\therefore E(\ln x_i)=\dfrac{1}{\theta}$，$E[\hat{g}(\theta)]=g(\theta)$

(3) $p(x;\theta)=\theta x^{\theta-1}$，$\ln p(x;\theta)=\ln\theta+(\theta-1)\ln x$

$$\frac{\partial\ln p(x;\ \theta)}{\partial\theta}=\frac{1}{\theta}+\ln x,\quad \frac{\partial^2\ln p(x,\ \theta)}{\partial\theta^2}=-\frac{1}{\theta^2}$$

$$I(\theta)=-E\left(-\frac{1}{\theta^2}\right)=\frac{1}{\theta^2}$$

(4) $\mathrm{Var}(\hat{g}(x))=\dfrac{1}{n^2}\sum\limits_{i=1}^{n}\mathrm{Var}(\ln x_i)=\dfrac{1}{n\theta^2}$，C-R 下界是 $[g'(\theta)]^2/nI(\theta)=\dfrac{1}{n\theta^2}$，故达到 C-R 下界。

2. (1) $p(x;\theta)=\dfrac{2\theta}{x^3}\mathrm{e}^{-\theta/x^2},x>0,\theta>0$

$$\ln p(x;\theta)=\ln(2\theta)-\frac{\theta}{x^2}-\ln x^3,\quad \frac{\partial\ln p(x,\theta)}{\partial\theta}=\frac{1}{\theta}-\frac{1}{x^2}$$

$$\frac{\partial^2\ln(x;\theta)}{\partial\theta^2}=-\frac{1}{\theta^2},\quad I(\theta)=-E\left(-\frac{1}{\theta^2}\right)=\frac{1}{\theta^2}$$

故 θ 无偏估计的 C-R 下界为$\dfrac{\theta^2}{n}$。

(2) $g(\theta)=1/\theta$ 的 C-R 下界为 $(g'(\theta))^2/nI(\theta)=\dfrac{1}{n\theta^2}$。

(3) $\because\ p(x_1,\cdots,x_n;\theta)=\dfrac{2^n\theta^n}{\prod\limits_{i=1}^{n}x_i^3}\mathrm{e}^{-\theta\sum\limits_{i=1}^{n}\frac{1}{x_i^2}}$

$$\ln p(x_1,\cdots,x_n;\theta)=n\ln 2+n\ln\theta-\theta\sum_{i=1}^{n}\frac{1}{x_i^2}-\ln\prod_{i=1}^{n}x_i^3$$

由$\dfrac{\partial\ln p}{\partial\theta}=0$ 可知，$1/\theta$ 的最大似然估计为 $\hat{g}(x)=\dfrac{1}{n}\sum\limits_{i=1}^{n}\dfrac{1}{x_i^2}$。

$E\left(\dfrac{1}{x^2}\right)=\dfrac{1}{\theta}$，故 $E[\hat{g}(x)]=\dfrac{1}{\theta}$，所以是无偏估计。

(4) $E\left(\dfrac{1}{x^2}\right)=\dfrac{2}{\theta^2}\Rightarrow\mathrm{Var}\left(\dfrac{1}{x}\right)=\dfrac{1}{\theta^2}$，所以 $\mathrm{Var}(\hat{g}(x))=\dfrac{1}{n\theta^2}$，故是有效估计。

3. （1）$\because p(x;\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$$\ln p(x,\sigma^2)=-\frac{(x-\mu)^2}{2\sigma^2}-\frac{1}{2}\ln(2\pi\sigma^2)$$

$$\frac{d\ln p(x,\sigma^2)}{d\sigma^2}=\frac{(x-\mu)^2}{2\sigma^4}-\frac{1}{2\sigma^2}$$

$$\frac{d^2\ln p(x,\sigma^2)}{d(\sigma^2)^2}=-\frac{(x-\mu)^2}{\sigma^6}+\frac{1}{2\sigma^4}$$

$$\therefore I(\sigma^2)=-E\left(-\frac{(x-\mu)^2}{2\sigma^6}+\frac{1}{2\sigma^4}\right)$$

$$=+\frac{1}{\sigma^6}E(x-\mu)^2-\frac{1}{2\sigma^4}$$

$$=\frac{1}{2\sigma^4}$$

可得 C-R 下界为$\frac{2\sigma^4}{n}$。

又$\because$ $\mathrm{Var}(\hat{\sigma}^2)=\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(x_i-\mu)^2$，$x_i-\mu$ 互相独立而$\left(\frac{x_i-\mu}{\sigma}\right)^2\sim\chi^2(1)$，则

$\mathrm{Var}(x_i-\mu)^2=2\sigma^4$

$\therefore \mathrm{Var}(\hat{\sigma}^2)=\frac{2\sigma^4}{n}$

故为有效估计。

（2）对于σ，$g(\sigma^2)=\sqrt{\sigma^2}$，$g'(\sigma^2)=\frac{1}{2\sigma}$，其 C-R 下界为$\frac{\sigma^2}{2n}$。

$$\because E\hat{\sigma}=\sqrt{\frac{\pi}{2}}\sum E|x_i-\mu|=\sqrt{\frac{\pi}{2}}\int_{-\infty}^{+\infty}|x-\mu|\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=\sigma\int_{0}^{+\infty}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}d(x-\mu)=\sigma$$

$\therefore$是无偏估计

又$\because$ $\mathrm{Var}(\hat{\sigma})=\frac{\pi}{2n^2}\sum\mathrm{Var}|x_i-\mu|=\frac{\pi}{2n^2}\sum[E(x_i-\mu)^2-(E|x_i-\mu|)^2]$

$$=\frac{\pi}{2n}\sigma^2$$

$\therefore$与 C-R 下界不相等，故不是有效估计。

4. （1）即求 $ET=\theta$。

$$ET=\sum_{j=1}^{4}a_jEN_j$$
$$=a_1EN_1+a_2EN_2+a_3EN_3+a_4EN_4$$
$$=a_1n(1-\theta)+a_2n(\theta-\theta^2)+a_3n(\theta^2-\theta^3)+a_4n\theta^3$$
$$=[a_1+(a_2-a_1)\theta+(a_3-a_2)\theta^2+(a_4-a_3)\theta^3]n$$

由 $ET=\theta$ 可得

$a_1=0$，$a_2=1/n$，$a_3=a_4=a_2=1/n$

即 $T=\dfrac{N_2+N_3+N_4}{n}$。

(2) $\mathrm{Var}(T)=\dfrac{\theta(1-\theta)}{n}$

$I(\theta)=\dfrac{n(1+\theta+\theta^2)}{\theta(1-\theta)}$，C-R 下界为$\dfrac{n(1-\theta)}{n(1+\theta+\theta^2)}$。

5. (1) $P(X)=\theta^n(1-\theta)^{\sum\limits_{i=1}^{n}x_i-n}, 0<\theta<1$

由因子分解定理知 $T=\sum\limits_{i=1}^{n}x_i$ 为 θ 的充分统计量。

$T=\sum\limits_{i=1}^{n}x_i$ 表示 n 次试验中成功 T 次。

$P(T=t)=\mathrm{C}_{n-1}^{t-1}\theta^n(1-\theta)^{t-n}$，$t=n$，$n+1$，…

为负二项分布。

对任一可积函数 $\varphi(t)$，$E_\theta\varphi(t)=0$

$$E\varphi(t)=\sum_{t=n}^{\infty}\mathrm{C}_{n-1}^{t-1}\varphi(t)\theta^n(1-\theta)^{t-n}=0$$

可得 $\varphi(t)=0$，说明 $T=\sum\limits_{i=1}^{\infty}x_i$ 是充分完备统计量。

(2) T 服从负二项分布，可知 $ET=\dfrac{n}{\theta}$，则有 $E\,\dfrac{T}{n}=\theta^{-1}$，即$\dfrac{T}{n}$是$\theta^{-1}$的无偏估计。

又因为 T 为 θ 的充分完备统计量，所以$\dfrac{T}{n}$是θ^{-1}的 UMVUE。

(3) $E\varphi(x_1)=P(x_1=1)=\theta$

所以 $\varphi(x_1)$ 是 θ 的无偏估计。

$$E[\varphi(x_1)\mid T=t]=\frac{P(x_1=1\mid T=t)}{P(T=t)}=\frac{\theta\begin{pmatrix}t-2\\n-2\end{pmatrix}\theta^{n-1}(1-\theta)^{t-n}}{\begin{pmatrix}t-1\\n-1\end{pmatrix}\theta^n(1-\theta)^{t-n}}=\frac{n-1}{t-1}$$

T 为 θ 的充分完备统计量，所以$\dfrac{n-1}{t-1}$为 θ 的 UMVUE。

6. x_i 来自正态总体 $N(\theta, 1)$，$p(x_i)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x_i-\theta)^2}{2}}$

$$P(X)=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{-\frac{\sum\limits_{i=1}^{n}(x_i-\theta)^2}{2}}=\frac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{-\frac{\sum\limits_{i=1}^{n}x_i^2-n\theta^2}{2}}$$

由因子分解定理知 $T=\sum_{i=1}^{n}x_i^2$ 为 θ^2 的充分统计量。

正态分布属于指数分布族，所以 $T=\sum_{i=1}^{n}x_i^2$ 是 θ^2 的充分完备统计量。

$$ET=\sum_{i=1}^{n}Ex_i^2=n(1+\theta^2)$$

所以 $E\left(\frac{T}{n}-1\right)=\theta^2$

即 $\frac{T}{n}-1$ 是 θ^2 的 UMVUE。

又因为 $p(X)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\theta)^2}{2}}$

$$\ln p(x)=-\ln\sqrt{2\pi}-\frac{(x-\theta)^2}{2}$$

$$\frac{\partial \ln p(x)}{\partial\theta}=x-\theta,\ \frac{\partial^2\ln p(x)}{\partial\theta^2}=-1$$

所以 $I(\theta)=-E(-1)=1$。

θ^2 的 C-R 下界为 $\frac{[(\theta^2)']^2}{nI(\theta)}=\frac{4\theta^2}{n}$。

因为 $\mathrm{Var}\left(\frac{T}{n}-1\right)=\frac{1}{n^2}\mathrm{Var}(T)=\frac{1}{n^2}\mathrm{Var}(\sum_{i=1}^{n}x_i^2)=\frac{\mathrm{Var}x_i^2}{n}$

$$=\frac{\mathrm{Var}[(x_i-\theta)^2+2(x_i-\theta)\theta+\theta^2]}{n}$$

$$=\frac{2+4\theta^2}{n}\neq\frac{4\theta^2}{n}$$

故不是有效估计。

习题 2.5

1. (1) $\pi(\theta)=1$，$\theta\in(0,1)$，$h(x,\theta)=\frac{2x}{\theta^2}$，$0<x<\theta<1$

$$m(x)=\int_x^1\frac{2x}{\theta^2}\mathrm{d}\theta=2(1-x)$$

$$\pi(\theta|x)=\frac{h(x,\theta)}{m(x)}=\frac{x}{\theta^2(1-x)},\ x<\theta<1$$

(2) $\pi(\theta)=3\theta^2$ $(0<\theta<1)$，$h(x,\theta)=6x$，$0<x<\theta<1$

$$m(x)=\int_x^1 6x\mathrm{d}\theta=6x(1-x)$$

$$\therefore\pi(\theta|x)=\frac{h(x,\theta)}{m(x)}=\frac{1}{1-x},\ x<\theta<1$$

2. 团体人的高度 $X\sim N(\theta,5^2)$，θ 的先验分布 $\pi(\theta)\sim N(172.72,2.54^2)$，样本均值为 176.53。故 θ 的后验分布为 $N(174.65,3.1746)$。

3. (1) 由题知 $\theta-\frac{1}{2}<12$，$12<\theta+\frac{1}{2}\Rightarrow 11.5<\theta<12.5$

$$h(x,\theta)=f(x\mid\theta)\pi(\theta)=\frac{1}{10}$$

$$\pi(\theta\mid x)=\frac{\frac{1}{10}}{\int_{11.5}^{12.5}\frac{1}{10}\mathrm{d}\theta}=1$$

故 θ 的后验分布为 $U(11.5,12.5)$。

(2) 6 个观察值 $x_{(1)}=10.9$，$x_{(6)}=11.7\Rightarrow 11.2<\theta<11.4$

$$\Rightarrow\pi(\theta\mid x)=\frac{\frac{1}{10}\cdot 1^6\cdot I\left\{\theta-\frac{1}{2}<x_{(1)}\right\}\cdot I\left\{\theta+\frac{1}{2}\geqslant x_{(n)}\right\}}{\int_{10}^{20}\frac{1}{10}\cdot 1^6\cdot I\left\{\theta-\frac{1}{2}\leqslant x_{(1)}\right\}\cdot I\left\{\theta+\frac{1}{2}\geqslant x_{(n)}\right\}\mathrm{d}\theta}$$

$$=\frac{\frac{1}{10}}{\int_{11.2}^{11.4}\frac{1}{10}\mathrm{d}\theta}$$

$\Rightarrow\theta$ 的后验分布为 $U(11.2,11.4)$

4. 泊松分布的概率函数为 $P(X=x\mid\lambda)=\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}$，若 λ 的先验分布为伽玛分布，其密度函数为：

$$\pi(\lambda)=\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}$$

对来自泊松分布 $P(\lambda)$ 的样本 x_1，x_2，…，x_n，λ 的后验分布为：

$$\pi(\lambda\mid x_1,x_2,\cdots,x_n)=\frac{\left(\prod_{i=1}^n\frac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}\right)\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}}{\int_0^{+\infty}\left(\prod_{i=1}^n\frac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}\right)\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}\mathrm{d}\lambda}$$

$$=\frac{\lambda^{\sum_{i=1}^n x_i+\alpha-1}\mathrm{e}^{-(\beta+n)\lambda}}{\int_0^{+\infty}\lambda^{\sum_{i=1}^n x_i+\alpha-1}\mathrm{e}^{-(\beta+n)\lambda}\mathrm{d}\lambda}$$

$$=\frac{(\beta+n)^{\sum_{i=1}^n x_i+\alpha}}{\Gamma\left(\sum_{i=1}^n x_i+\alpha\right)}\lambda^{\sum_{i=1}^n x_i+\alpha-1}\mathrm{e}^{-(\beta+n)\lambda}$$

即 λ 的后验分布为 $\Gamma\left(\sum_{i=1}^n x_i+\alpha,\beta+n\right)$，仍为伽玛分布，即证。

5. x_1，x_2，…，x_n 与 θ 的联合分布为：

$$h(x_1,\cdots,x_n,\theta)=\frac{1}{\theta^n}\frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}，\ \theta>\theta_0,x_{(n)}<\theta$$

要使 $\theta>\theta_0$ 与 $\theta>x_{(n)}$ 同时成立，必须 $\theta>\max(x_{(n)}, \theta_0)$，所以 θ 的后验分布为：

$$\pi(\theta \mid x_1,\cdots,x_n)=\frac{\frac{1}{\theta^n}\frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}}{\int_{\max(x_{(n)},\theta_0)}^{+\infty}\frac{1}{\theta^n}\frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}}\mathrm{d}\theta}=\frac{\frac{1}{\theta^{n+\alpha+1}}}{\int_{\max(x_{(n)},\theta_0)}^{+\infty}\frac{1}{\theta^{n+\alpha+1}}\mathrm{d}\theta}$$

$$=\frac{(n+\alpha)[\max(x_{(n)},\theta_0)]^{n+\alpha}}{\theta^{n+\alpha+1}},\ \theta>\max(x_{(n)},\theta_0)$$

这是一个参数为 $n+\alpha$ 与 $\max(x_{(n)}, \theta_0)$ 的 Pareto 分布。

6. $P(\boldsymbol{x}|\theta)=\frac{1}{\theta^3}I\{\theta\geqslant 8\}$

$\pi(\theta)=\frac{192}{\theta^4}I\{\theta\geqslant 4\}$

$h(x,\theta)=\frac{192}{\theta^7}I\{\theta\geqslant 8\}$

$m(x)=\int_8^{\infty}\frac{192}{\theta^7}\mathrm{d}\theta=\frac{1}{8\,192}$

$\therefore \pi(\theta|x)=\frac{h(x,\theta)}{m(x)}=\frac{1\,572\,864}{\theta^7}I\{\theta\geqslant 8\}$

7. 样本和 θ 的联合密度函数为 $h(x_1,\cdots,x_n;\theta)=\theta^n(1-\theta)^{\sum_{i=1}^n x_i}$，于是

$$\pi(\theta \mid x_1,\cdots,x_n)=\frac{h(x_1,\cdots,x_n,\theta)}{\int h(x_1,\cdots,x_n,\theta)\mathrm{d}\theta}=\frac{\theta^n(1-\theta)^{\sum_{i=1}^n x_i}}{\int_0^1\theta^n(1-\theta)^{\sum_{i=1}^n x_i}\mathrm{d}\theta}$$

$$=\frac{\Gamma\left(n+\sum_{i=1}^n x_i+2\right)}{\Gamma(n+1)\Gamma\left(\sum_{i=1}^n x_i+1\right)}\theta^n(1-\theta)^{\sum_{i=1}^n x_i}$$

因此 θ 的后验分布是 $Be(n+1,\ \sum_{i=1}^n x_i+1)$

作 1 次观察，观察值为 3，则 θ 的贝叶斯估计为 $\hat{\theta}_B=\frac{2}{4+2}=\frac{1}{3}$，$\theta\sim Be(2,\ 4)$。

作 3 次观察，观察值为 2，3，5，$\theta\sim Be(4,\ 11)$ 则 $\hat{\theta}_B=\frac{4}{11+4}=\frac{4}{15}$。

8. (1) $\pi(\theta \mid x)\propto P(X=k \mid \theta)\pi(\theta)$

$=\theta(1-\theta)^k\pi(\theta)$

$\pi(\theta \mid X)$ 与 $\pi(\theta)$ 同分布，易知服从 $Be(a,\ b)$。

(2) $\pi(\theta)$ 服从 $Be(a,\ b)$。

则 $\pi(\theta \mid X)\sim Be(a+1,\ b+k)$

$\therefore E(\theta|X)=\frac{a+1}{a+b+k+1}$

$$\mathrm{Var}(\theta|X)=\frac{(a+1)(b+k)}{(a+b+k+1)^2(a+b+k+2)}$$

9. $\because\lambda$ 服从伽玛分布 $Ga(\alpha,\ \beta)$

$$\therefore\begin{cases}\dfrac{\alpha}{\beta}=0.2\\[2ex]\dfrac{\alpha}{\beta^2}=1\end{cases}\Rightarrow\begin{cases}\alpha=0.04\\\beta=0.2\end{cases}$$

即 λ 的先验分布为 $Ga(0.04,\ 0.2)$。

$$\therefore\pi(\lambda)=\frac{0.2^{0.04}}{\Gamma(0.04)}\lambda^{0.04-1}\mathrm{e}^{-0.2\lambda}$$

$$\pi(\lambda\mid\boldsymbol{x})=Ga(20.04,\ 60.2)$$

$$\therefore\hat{\lambda}_B=\frac{20.04}{60.2}=0.333$$

又 $\because\pi\left(\dfrac{1}{\lambda}\mid\boldsymbol{x}\right)=IGa(20.04,\ 60.2)$

$$\therefore\hat{\theta}_B=\frac{60.2}{19.04}=3.16$$

10. $p(x_i)=\dfrac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}$

$$p(\boldsymbol{x})=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}\mathrm{e}^{-n\lambda}$$

$$\pi(\lambda)\sim Ga(3,1)$$

$$\therefore\pi(\lambda\mid x)\propto\lambda^{\sum_{i=1}^{n}x_i}\mathrm{e}^{-n\lambda}\lambda^2\mathrm{e}^{-\lambda}$$

$$=\lambda^{\sum_{i=1}^{n}x_i+2}\mathrm{e}^{-(n+1)\lambda}$$

$$\Rightarrow\pi(\lambda\mid x)\sim Ga\left(\sum_{i=1}^{n}x_i+3,n+1\right)$$

题中 $n=3$，$\sum_{i=1}^{n}x_i=8$，故 λ 的贝叶斯估计 $\hat{\lambda}=\dfrac{11}{4}$。

习题 3.1

1. (1) $n=4$，$\dfrac{\bar{x}-\mu}{\sqrt{\frac{1}{4}}}=2(\bar{x}-\mu)\sim N(0,\ 1)$

$$P(\bar{x}-1\leqslant\mu\leqslant\bar{x}+1)=P(-2\leqslant2(\bar{x}-\mu)\leqslant2)$$
$$=2\Phi(2)-1=0.9544$$

$\therefore$ 置信系数是 0.954 4。

(2) $\sqrt{n}(\bar{x}-\mu)\sim N(0,\ 1)$

$$P(\bar{x}-1\leqslant\mu\leqslant\bar{x}+1)=P(-\sqrt{n}\leqslant\sqrt{n}(\bar{x}-\mu)\leqslant\sqrt{n})$$

$$=2\Phi(\sqrt{n})-1\geqslant 0.99\Rightarrow n\geqslant 6.63$$

∴ 至少需要的样本量为 7。

2. 由例 3.1.3 可知

$$P\left(\frac{2T_n}{\chi^2_{1-\alpha/2}(2n)}\leqslant\frac{1}{\lambda}\leqslant\frac{2T_n}{\chi^2_{\alpha/2}(2n)}\right)=1-\alpha$$

$n=1, 2\chi^2_{\alpha/2}(2)=\chi^2_{1-\alpha/2}(2)\Rightarrow\alpha=0.2387$

3. (1) 令 $Y=n(x_{(1)}-\theta)$

$$P(Y\leqslant y)=P[n(x_{(1)}-\theta)\leqslant y]=P\left(x_{(1)}\leqslant\frac{y}{n}+\theta\right)$$

$$=\begin{cases}1-e^{-y}, & y\geqslant 0\\ 0, & y<0\end{cases}$$

∴$n(x_{(1)}-\theta)\sim\exp(1)$，分布与 θ 无关，故是枢轴量。

(2) ①$p(y)=e^{-y}(y\geqslant 0)$　递减，$P(c\leqslant n(x_{(1)}-\theta)\leqslant d)=P(x_{(1)}-\frac{d}{n}\leqslant\theta\leqslant x_{(1)}-\frac{c}{n})$

∴当 $d-c$ 最小时区间最短，p（y）递减

∴当 $c=0$　$d=-\ln\alpha$ 时区间最短

∴置信区间是$\left[x_{(1)}+\frac{\ln\alpha}{n}, x_{(1)}\right]$

②$p(c\leqslant n(x_{(1)}-\theta)\leqslant d)=1-\alpha$

求 χ^2 的$\frac{\alpha}{2}$与 $1-\frac{\alpha}{2}$分位数

可得 $c=x_{(1)}-x^2_{1-\frac{\alpha}{2}}(2)/n$

$d=x_{(1)}-x^2_{\frac{\alpha}{2}}(2)/n$

4. (1) $F(x;\theta)=1-\frac{\theta}{x}$，$0<\theta<x<\infty$

$\frac{x_{(1)}}{\theta}$的分布函数 $F_1(x, \theta)=1-\frac{1}{x^n}$

$$P\left(c\leqslant\frac{x_{(1)}}{\theta}\leqslant d\right)=1-\frac{1}{d^n}-\left(1-\frac{1}{c^n}\right)=\frac{1}{c^n}-\frac{1}{d^n}=1-\alpha$$

即　$P\left(\frac{x_{(1)}}{d}\leqslant\theta\leqslant\frac{x_{(1)}}{c}\right)=1-\alpha$

置信区间为$\left[\frac{x_{(1)}}{d}, \frac{x_{(1)}}{c}\right]$。

(2) $P(a\leqslant\theta\leqslant b)\Rightarrow P\left(\frac{x_{(1)}}{b}\leqslant\frac{x_{(1)}}{\theta}\leqslant\frac{x_{(1)}}{a}\right)$

$$=1-\left(\frac{a}{x_{(1)}}\right)^n-\left[1-\left(\frac{b}{x_{(1)}}\right)^n\right]$$

$$=\left(\frac{b}{x_{(1)}}\right)^n-\left(\frac{a}{x_{(1)}}\right)^n=1-\alpha$$

等式两边对 b 求导 $\left(\frac{b}{x_{(1)}}\right)^{n-1}-\left(\frac{a}{x_{(1)}}\right)^{n-1}\frac{\partial a}{\partial b}=0$

∴$\frac{\partial a}{\partial b}=\left[\frac{b}{a}\right]^{n-1}>1$

$\therefore \frac{\partial(b-a)}{\partial b}=1-\frac{\partial a}{\partial b}<0$

当 b 取最大值 $x_{(1)}$ 时，$b-a$ 最小。代入得 $a=x_{(1)}\sqrt[n]{\alpha}$，$b=x_{(1)}$。即区间长度最短的置信区间为 $[x_{(1)}\sqrt[n]{\alpha}, x_{(1)}]$。

5. 由例 3.1.3 可知：

(1) $[2T_n/\chi^2_{1-\alpha/2}(2n), 2T_n/\chi^2_{\alpha/2}(2n)]$；

(2) $[(2T_n/\chi^2_{1-\alpha/2}(2n))^2, (2T_n/\chi^2_{\alpha/2}(2n))^2]$；

(3) $[e^{-\chi^2_{1-\alpha/2}(2n)/2T_n}, e^{-\chi^2_{\alpha/2}(2n)/2T_n}]$。

6. (1) $P(x_{(1)}<x_{0.5}<x_{(n)})$

$$
\begin{aligned}
&=1-P(x_{(1)}\geqslant x_{0.5})-P(x_{0.5}\geqslant x_{(n)})\\
&=P(x_{(1)}<x_{0.5})-P(x_{(n)}\leqslant x_{0.5})\\
&=1-[1-0.5]^n-0.5^n\\
&=1-\left(\frac{1}{2}\right)^{n-1}=1-0.5^{n-1}
\end{aligned}
$$

(2) 见 (3)，$1-0.5^{n-1}\times(n+1)$

(3) $P(x_{(k)}<x_{0.5}<x_{(n-k+1)})$

$$
\begin{aligned}
&=P(x_{(k)}<x_{0.5})-P(x_{(n-k+1)}\leqslant x_{0.5})\\
&=\int_0^{\frac{1}{2}}\frac{n!}{(k-1)!(n-1)!}[y^{k-1}(1-y)^{n-(k-1)-1}-y^{n-(k-1)-1}(1-y)^{k-1}]\mathrm{d}y
\end{aligned}
$$

(由 $x_{(k)}$ 的密度函数)

$$
\begin{aligned}
&=\int_0^{\frac{1}{2}}\frac{n!}{(k-1)!(n-k)!}y^{k-1}(1-y)^{n-(k-1)-1}\mathrm{d}y-\int_1^{\frac{1}{2}}\frac{n!}{(k-1)!(n-k)!}(1-y)^{n-(k-1)-1}y^{k-1}\mathrm{d}(-y)\\
&=\int_0^{\frac{1}{2}}\frac{n!}{(k-1)!(n-k)!}y^{k-1}(1-y)^{n-(k-1)-1}\mathrm{d}y\\
&\quad-\int_{\frac{1}{2}}^{1}\frac{n!}{(k-1)!(n-k)!}(1-y)^{n-(k-1)-1}y^{k-1}\mathrm{d}y\\
&=1-2\int_{\frac{1}{2}}^{1}\frac{n!}{(k-1)!(n-k)!}(1-y)^{n-k}y^{k-1}\mathrm{d}y
\end{aligned}
$$

因为 $\sum_{k=0}^{r}\binom{n}{k}p^k(1-p)^{n-k}=\frac{n!}{r!(n-r)!}\int_p^1 x^r(1-x)^{n-r-1}\mathrm{d}x$

$$
\begin{aligned}
\therefore P(x_{(k)}<x_{0.5}<x_{(n-k+1)})&=1-2\times 0.5^n\left[\sum_{i=0}^{k-1}\binom{n}{i}\right]\\
&=1-\left(\frac{1}{2}\right)^{n-1}(C_n^0+C_n^1+\cdots+C_n^{k-1})
\end{aligned}
$$

习题 3.2

1. 标准差 $\sigma=1.19$ 已知，故均值 μ 的置信区间为：

$[\bar{x}-\mu_{1-\alpha/2}\sigma/\sqrt{n}, \bar{x}+u_{1-\alpha/2}\sigma/\sqrt{n}]$

所以该化纤强力均值 μ 的置信区间水平为 0.95 的置信区间为：

[5.77，6.93]

2. 总体方差 σ^2 未知，样本标准准差 $s=6$，所以置信水平为 0.95 的置信区间为 $[\bar{x}\pm t_{1-\alpha/2}(n-1)s/\sqrt{n}]$，

代入数据可知置信区间 [23.29，36.71]。

3. σ 未知，故 μ 的置信区间为 $[\bar{x}\pm t_{1-\alpha/2}(n-1)s/\sqrt{n}]$

∴ 为 [54.74，75.54]

4. σ 未知，故 μ 的置信区间为 $[\bar{x}\pm t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}}]$

∴ 为 [2.63，3.97]

5. $\bar{x}=6$，$s=\sqrt{\frac{1}{8}\sum_{i=1}^{9}(x_i-\bar{x})^2}=0.57$，$t_{0.95}(8)=1.86$

∴ $\bar{x}+t_{0.95}(8)\frac{s}{\sqrt{9}}=6.353$

6. 总体方差 σ^2 未知，s 已知，其置信下限为 $\bar{x}-t_{0.95}\frac{s}{\sqrt{n}}$，代入得 294.9。

7. $x\sim LN(\mu,1)$，则 $Y=e^x\sim N(\mu,1)$。

置信上限为：$\bar{Y}+u_{0.9}\frac{s}{\sqrt{4}}$

代入数据可得其置信上限为 5.372 45。

8. 由 3.2.3 节内容可知，σ^2 的 $1-\alpha$ 置信区间为：

$$\left[\frac{\sum x_i^2}{\chi^2_{1-\alpha/2(n)}},\frac{\sum x_i^2}{\chi^2_{\alpha/2(n)}}\right]$$

9. 由 3.2.2 节内容可知，标准差 σ 已知场合

$$n\geqslant\left(\frac{u_{1-\alpha/2}\sigma}{d}\right)^2$$

代入可得至少要测量 15 次。

10. 标准差 σ 未知场合

$$n\geqslant\left(\frac{t_{1-\alpha/2}(n_0-1)s_0}{d}\right)^2$$

代入可得至少应取 12 个

11. 由例 3.2.5 可知

$$n=\left(\frac{0.84\times t_{0.975}(7)}{0.5}\right)^2$$

$$n_1=n-8$$

查表可得

$$n_1=8$$

12. 两样本源自方差相等的正态总体，其置信区间为：

$$\bar{x}-\bar{y}\pm t_{1-\alpha/2}(n_1+n_2-2)\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)s_w}$$

$$s_w=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

代入数据可得 0.95 的置信区间为 [−0.39，0.49]。

13. (1) 由 4.5.2 节内容可知，$\frac{\sigma_1^2}{\sigma_2^2}$的 $1-\alpha$ 置信区间为：

$$\left[\frac{s_1^2}{s_2^2}\frac{1}{F_{1-\alpha/2}(n_1-1,\ n_2-1)},\ \frac{s_1^2}{s_2^2}F_{1-\alpha/2}(n_2-1,\ n_1-1)\right]$$

(2) $\frac{\sigma_1}{\sigma_2}$的 $1-\alpha$ 单侧置信下限：

$$\frac{s_1}{s_2}\sqrt{F_\alpha(n_2-1,\ n_1-1)}$$

14. $\frac{Q}{\sigma^2}\sim\chi^2(n-1)$

$$\therefore \frac{\sum_{i=1}^{5}Q_i}{\sigma^2}\sim\chi^2(n-5),\ n=\sum_{i=1}^{5}n_i$$

$\therefore$ 置信区间为$\left[\frac{\sum_{i=1}^{5}Q_i}{\chi^2_{1-\alpha/2}(n-5)},\ \frac{\sum_{i=1}^{5}Q_i}{\chi^2_{\alpha/2}(n-5)}\right]$，代入得 [4.78，15.58]。

习题 3.3

1. $n=100$，$\bar{x}=0.16$，$\alpha=0.05$，$u_{0.975}=1.96$

由例 3.3.4 可知

$a=100+1.96^2$

$b=-(2\times100\times0.16+1.96^2)$

$c=100\times0.16^2$

$\therefore \hat{P}_L=0.1$，$\hat{P}_U=0.24$

即置信区间为 [0.1，0.24]。

2. $\hat{p}=0.65$

其置信区间为 $\hat{p}\pm u_{0.975}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$，即 [0.62，0.68]。

3. 步骤同第 2 题，得置信区间为 [0.496，0.624]。

4. $Ex=\lambda$，$Dx=\lambda$，由例 3.3.4 可知，$\frac{\bar{x}-\lambda}{\sqrt{\lambda/n}}\dot\sim N(0,1)$

$$P\left(\left|\frac{\bar{x}-\lambda}{\sqrt{\lambda/n}}\right|\leqslant u_{1-\alpha/2}\right)=1-\alpha$$

即 $(\bar{x}-\lambda)^2\leqslant u_{1-\alpha/2}^2\frac{\lambda}{n}$

即 $n\lambda^2-(2n\bar{x}+u_{1-\alpha/2}^2)\lambda+n\bar{x}^2\leqslant 0$

求出两个实根，可得其 $1-\alpha$ 的置信区间为：

$$(\bar{x}+u_{1-\alpha/2}^2/(2n))\pm\frac{1}{2}\sqrt{(2\bar{x}+u_{1-\alpha/2}^2/n^2)-4\bar{x}^2}$$

5. $\bar{x}=1.8$，$\bar{x}-u_{1-\alpha/2}\sqrt{\frac{1.8}{n}}=1.636$

6. 其置信区间为 $\bar{x}\pm u_{1-\alpha/2}\sqrt{\frac{\lambda}{n}}$，可得 $[0.715\,3，0.777\,3]$。

7. 如例 3.3.5

$$n\geqslant\left(\frac{u_{0.995}}{2\times 0.001}\right)^2=1\,658\,944$$

习题 3.4

1. 由例 2.5.5 可知，θ 的后验分布为：

$$N\left(3，\frac{1}{4}\right)$$

∴ θ 的 0.95 可信区间为：

$$[2.02，3.98]$$

2. $p(\lambda|x)\propto\ \lambda^{a-1}\mathrm{e}^{-(n+b)\lambda}$

即 θ 的后验分布为：

$$Ga(a，b+n)$$

可得可信区间为：

$$\left[\frac{\chi_{\alpha/2}^2(2a)}{2(b+n)},\frac{\chi_{1-\frac{\alpha}{2}}^2(2a)}{2(b+n)}\right]$$

（详见例 3.4.2。）

3. （1）$p(\boldsymbol{x}|\theta)=\frac{1}{\theta^n}I\{\theta\geqslant x_{(n)}\}$

$$\pi(\theta)=\frac{\beta\theta_0^\beta}{\theta^{\beta+1}}I(\theta>\theta_0)$$

$$\pi(\theta|\boldsymbol{x})=\frac{(\beta+n)[\max\{\theta_0，x_{(n)}\}]^{\beta+n}}{\theta^{\beta+n+1}}$$

$$\therefore\ \hat{\theta}_B=E(\theta|\boldsymbol{x})=\frac{(\beta+n)\max\{\theta_0，x_{(n)}\}}{(\beta+n-1)}$$

（2）$P\{\theta\leqslant c\}=1-\alpha$，$\pi(\theta|\boldsymbol{x})$ 递减，又 θ 的后验分布函数为：

$$F(\theta|\boldsymbol{x})=1-\left(\frac{\max(\theta_0，x_{(n)})}{\theta}\right)^{\beta+n}$$

令 $F(c|\boldsymbol{x})=1-\alpha$

得 $c=\max\{\theta_0,x_{(n)}\}\alpha^{-\frac{1}{\beta+n}}$

即 θ 的 $1-\alpha$ 可信上限为 $\max\{\theta_0，x_{(n)}\}\alpha^{-\frac{1}{\beta+n}}$。

4. $p(\sigma^2 \mid \boldsymbol{x}) \propto \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \mathrm{e}^{-\frac{\sum\limits_{i=1}^n x_i^2}{2\sigma^2}} \frac{\lambda^\alpha}{\Gamma(\alpha)(\sigma^2)^{\alpha+1}} \mathrm{e}^{-\frac{\lambda}{\sigma^2}}$

$$\propto \frac{\lambda^\alpha}{\Gamma(\alpha)(\sigma^2)^{\frac{n}{2}+\alpha+1}} \mathrm{e}^{-\frac{\sum\limits_{i=1}^n x_i^2+2\lambda}{2\sigma^2}}$$

$$\propto \frac{1}{(\sigma^2)^{\frac{n}{2}+\alpha+1}} \mathrm{e}^{-\frac{(\sum\limits_{i=1}^n x_i^2+2\lambda)/2}{\sigma^2}}$$

即 σ^2 的后验分布为 $IGa\left(\frac{n}{2}+\alpha, \frac{\sum\limits_{i=1}^n x_i^2+2\lambda}{2}\right)$。

由例 3.4.2 可得其 $1-\alpha$ 可信区间为：

$$\left[\frac{2\lambda+\sum\limits_{i=1}^n x_i^2}{\chi^2_{1-\alpha/2}(2\alpha+n)}, \frac{2\lambda+\sum\limits_{i=1}^n x_i^2}{\chi^2_{\alpha/2}(2\alpha+n)}\right]$$

5. $p(\theta \mid \boldsymbol{x}) \propto \frac{1}{(\sqrt{2\pi})^n} \mathrm{e}^{-\frac{\sum\limits_{i=1}^n (x_i-\theta)^2}{2}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(\theta-x_0)^2}{2}}$

$$\propto \mathrm{e}^{-\frac{\sum\limits_{i=0}^n (x_i-\theta)^2}{2}}$$

$$\propto \mathrm{e}^{-\frac{\left(\theta-\frac{\sum\limits_{i=0}^n x_i}{n+1}\right)^2}{2\frac{1}{n+1}}}$$

即 θ 的后验分布为 $N\left(\frac{\sum\limits_{i=0}^n x_i}{n+1}, \frac{1}{n+1}\right)$，所以 $1-\alpha$ 可信区间为 $\frac{\sum\limits_{i=1}^n x_i}{n+1} \pm u_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{n+1}}$。

6. $p(\theta \mid \boldsymbol{x}) = \theta^n (x_1 \cdots x_n)^{\theta-1} \theta^{\gamma-1} \mathrm{e}^{-\lambda\theta}$

$$\propto \theta^{n+\gamma-1} \mathrm{e}^{\sum\limits_{i=1}^n \ln x_i \theta - \lambda\theta}$$

$\therefore p(\theta \mid \boldsymbol{x}) \sim Ga(n+\gamma, \lambda - \sum\limits_{i=1}^n \ln x_i)$

$\therefore \theta$ 的 $1-\alpha$ 等尾可信区间为：

$$\left[\frac{\chi^2_{\alpha/2}(2n+2\gamma)}{2(\lambda-\sum\limits_{i=1}^n \ln x_i)}, \frac{\chi^2_{1-\alpha/2}(2n+2\gamma)}{2(\lambda-\sum\limits_{i=1}^n \ln x_i)}\right]$$

习题 4.1

1. 统计假设的对象是总体的未知参数，故（1）（3）（4）（7）为合理的统计假设。
2. （1）（4）（6）（7）（8）。
3. （1）H_0：$\theta=60$ vs H_1：$\theta<60$。

（2）该检验为单侧检验，且 w 的方向应和 H_1 的方向相同，故为 W_2。

(3) $W=\left\{\bar{x}\leqslant u_0-u_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right\}$，即 $\{\bar{x}\leqslant 54.85\}$。

(4) 即求 $\alpha=\Phi(-2.33)\Rightarrow\alpha=0.01$

4. $H_0:\mu=50\quad vs\quad H_1:\mu\neq 50$

$$u_{1-\alpha/2}=1.96$$

$$u_0=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{5}{3}<1.96$$

不能拒绝原假设，故正常。

5. (1) $W=\{|u|\geqslant c\}$，$\alpha=0.05$，故 $c=u_{1-\alpha/2}=1.96$，即 $W=\{|u|\geqslant 1.96\}$。

(2) $\bar{x}=101.2\Rightarrow u=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{101.2-100}{2/\sqrt{9}}=1.8<1.96$，故不能拒绝原假设。

(3) $\beta=P_\mu\left(-c<\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}<c\right)$

$$=P_\mu\left(-c<\frac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}+\frac{\mu-\mu_0}{\sigma_0/\sqrt{n}}<c\right)=\Phi\left(c+\frac{\mu_0-\mu}{\sigma_0/\sqrt{n}}\right)-\Phi\left(-c+\frac{\mu_0-\mu}{\sigma_0/\sqrt{n}}\right)$$

代入数据得

$$\beta=0.005\,5$$

6. (1) $x_{(n)}$的分布函数为：

$$F(y)=\left(\frac{y}{\theta}\right)^n$$

在 H_0 为真的情况下

$$P\{x_{(n)}\leqslant 1.5\}=0.75^n$$

(2) $0.75^n\leqslant 0.05\Rightarrow n\geqslant 11$

7. $T\sim b(n, p)$

$$g(p)=P_p(T\geqslant 8)=\sum_{j=8}^{20}\binom{20}{8}p^8(1-p)^{12}=\frac{\Gamma(21)}{\Gamma(8)\Gamma(13)}\int_0^p u^7(1-u)^{12}\mathrm{d}u$$

$\alpha=g(0.2)=0.032\,1$，$\beta=1-g(0.4)=0.415\,9$

8. (1) 第Ⅱ类错误；(2) 第Ⅰ类错误。

9. (1) $H_0:\theta\geqslant 110\quad vs\quad H_1:\theta<110$

$$u_0=\frac{\bar{x}-110}{\sigma/\sqrt{n}}=\frac{108.2-110}{4/5}=-2.25<u_{0.05}=-1.64$$

拒绝原假设，不正常。

(2) $g(u)=\Phi\left(u_\alpha-\frac{\mu-110}{0.8}\right)=1-\Phi\left(u_{1-\alpha}+\frac{\mu-110}{0.8}\right)$，图略。

(3) $\beta(108)=0.196\,3$。

10. $g(\theta)=\begin{cases} P\left(\dfrac{\overline{x}-2}{1/\sqrt{n}}\geqslant\dfrac{2.6-2}{1/\sqrt{n}}\right), & \mu=2 \\ P\left(\dfrac{\overline{x}-3}{1/\sqrt{n}}\geqslant\dfrac{2.6-3}{1/\sqrt{n}}\right), & \mu=3 \end{cases}$

当 $n\to\infty$ 时

$g(\theta)=\begin{cases} 1-\Phi(\infty), & \mu=2 \\ 1-\Phi(-\infty), & \mu=3 \end{cases}$

$\therefore \alpha=1-\Phi(\infty)=0$

$\beta=1-g(\theta)=\Phi(-\infty)=0$

习题 4.2

1. $H_0:\mu=500 \quad \text{vs} \quad H_1:\mu\neq 500$

$u_0=\dfrac{496-500}{\sigma/\sqrt{n}}=\dfrac{-4}{5/5}=-4<u_{\alpha/2}=-1.96$

拒绝原假设。

故不正常。

2. $H_0:\mu\leqslant 6 \quad \text{vs} \quad H_1:\mu>6$

$u_0=\dfrac{\overline{x}-\mu}{\sigma/\sqrt{n}}=\dfrac{6.35-6}{1.19/\sqrt{100}}>u_{1-\alpha}=1.64$

拒绝原假设，即工艺改进。

3. $H_0:\mu\geqslant 100 \quad \text{vs} \quad H_1:\mu<100$

$u_0=\dfrac{\overline{x}-\mu}{\sigma/\sqrt{n}}=-3$

$P=P(u\leqslant -3)=0.00135$ 很小，拒绝 H_0，即平均成本有所下降.

4. $H_0:\mu=0.618 \quad \text{vs} \quad H_1:\mu\neq 0.618$

由题可得出 20 个样本的均值与方差，用统计量 $t_0=\dfrac{\overline{x}-\mu}{s/\sqrt{n}}$进行检验。通过计算可得出不能拒绝原假设，均值可以为 0.618。

5. $H_0:\mu=0 \quad \text{vs} \quad H_1:\mu>0$

$t_0=\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}=10.88>t_{0.95}(49)$，拒绝 H_0，即该中药对治理高血压有效。

6. $H_0:\mu\geqslant 950 \quad \text{vs} \quad H_1:\mu<950$

$u_0=\dfrac{\overline{x}-950}{\sigma/\sqrt{n}}=-6.6$

$P\{u\leqslant u_0\}=0$，故拒绝 H_0，即认为存在显著降低。

7. $P\{u\leqslant u_0\}=\Phi(-1.25)=0.1056$

8. $P\{u>2.94)=1-\Phi(2.94)=0.0016$

9. $P(t(9)\leqslant -3)=0.008$

10. 步骤见例 4.2.8，可知

(1) $n=\frac{(u_{1-\alpha}+u_{1-\beta})^2\sigma^2}{(\mu_1-\mu_0)^2}=7$

(2) $n\approx\frac{(u_{1-\alpha/2}+u_{1-\beta})^2\sigma^2}{(\mu_1-\mu_0)^2}=8$

习题 4.3

1. $H_0:\mu_1\leqslant\mu_2$，$H_1:\mu_1>\mu_2$

$$u=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}=0.834$$

$$W=\{u\geqslant u_{1-\alpha}\}=\{u\geqslant 1.645\}$$

所以不能拒绝 H_0，无显著提高。

2. $\bar{x}_1=16.015$，$\bar{x}_2=16.005$

$$u_0=\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=0.9877$$

(1) u_0 较小，接受 H_0；

(2) $P(u\leqslant u_0)=\Phi(0.9877)$

$=0.8384$

(3) $\mu_1-\mu_2$ 的置信区间为 $(\bar{x}_1-\bar{x}_2)\pm u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$，代入可得 $[-0.0098, 0.0298]$。

(4) $n=\frac{(u_{1-\alpha}+u_{1-\beta})^2(\sigma_1^2+\sigma_2^2)}{(\mu_1-\mu_2)^2}$，代入数据可得 $n\geqslant 12$。

3. (1) 两样本的方差已知

$$u=\frac{\overline{X}-\overline{Y}-10}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}$$

拒绝域为：

$$W=\{u\geqslant u_{1-\alpha}\}$$

(2) $u_0=\frac{162.7-151.8-10}{\sqrt{\frac{1}{10}+\frac{1}{12}}}=2.102$

$u_{1-\alpha}=1.645$，$u_0>u_{1-\alpha}$，故拒绝 H_0。

(3) $\mu_1-\mu_2$ 的 0.95 单侧置信下限为：

$$\overline{X}-\overline{Y}-u_{1-\alpha}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}=10.19$$

4. H_0：$\mu_1-\mu_2\leqslant 0$ vs H_1：$\mu_1-\mu_2>0$

$\overline{X}_1=0.0273$，$s_1=0.1677$，$\overline{X}_2=0.1327$，$s_2=0.0801$

处理前与处理后的方差是否相等，我们并不可知，故采取统计量

$$t^* = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t(l)$$

$$l = \left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 \Big/ \left(\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}\right)$$

$$\alpha = 0.05$$

代入数据，拒绝原假设，即显著降低了含脂率。

5. $H_0: |\mu_1 - \mu_2| = 5$ vs $H_1: |\mu_1 - \mu_2| \neq 5$

$$u = \frac{|\overline{X} - \overline{Y}| - |\mu_1 - \mu_2|}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = -5.14$$

$$W = \{|u| \geqslant u_{1-\alpha/2}\} = \{|u| \geqslant 1.96\}$$

故拒绝原假设，没有达到设计要求。

6. (1) $H_0: \mu_1 - \mu_2 = 0$ vs $H_1: \mu_1 - \mu_2 \neq 0$

方差相等且未知，则统计量

$$t = \frac{\overline{X} - \overline{Y}}{s_w \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t(n+m-2)$$

$$s_w^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{m+n-2}$$

代入数据

$$t_0 = 0.23 < t_{1-\alpha/2} = 2.042$$

故不能拒绝 H_0，即直径相等。

(2) $P(|t| < 0.23) = 0.82$。

(3) $\mu_1 - \mu_2$ 的置信区间为：

$$(\overline{X} - \overline{Y}) \pm t_{1-\alpha/2}(m+n-2) s_w \sqrt{\frac{1}{n} + \frac{1}{m}}$$

即 $[-0.3940, 0.4940]$。

7. 同第6题，求得90%的置信区间为 $[-3.216, -2.286]$，不含0，故 μ_1 与 μ_2 间有显著差异。

8. (1) $H_0: \mu_1 - \mu_2 = 0$ vs $H_1: \mu_1 - \mu_2 \neq 0$

$$t = \frac{\overline{X} - \overline{Y}}{s_w \sqrt{\frac{1}{n} + \frac{1}{m}}} = -2.5452$$

$$W = \{|t| \geqslant t_{1-\alpha/2}(m+n-2)\} = \{|t| \geqslant 2.101\}$$

拒绝 H_0，即蚀刻率不相等。

(2) $P\{|t| \geqslant t_0\} = 0.012$。

(3) $\overline{X} - \overline{Y} \pm t_{1-\alpha/2}(m+n-2) s_w \sqrt{\frac{1}{n} + \frac{1}{m}}$

即　$[-0.6937, -0.0663]$。

（4）略。

习题 4.4

1. $H_0: \mu_d \leqslant 0$ vs $H_1: \mu_d > 0$，$\mu_d = \mu_{后} - \mu_{前}$

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{3.585714}{5.271125/\sqrt{7}} = 1.7998$$

$$W = \{t > t_{1-\alpha}(n-1)\},\ t_{0.95}(6) = 1.943$$

接受 H_0，即不能认为该道工序对提高参数值有用。

2. $H_0: \mu_d \geqslant 0$ vs $H_1: \mu_d < 0$，$\mu_d = \mu_x - \mu_y$

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} = -3.867 < t_{0.05}(6) = -1.943$$

所以应拒绝 H_0，即平均打字速度提高了。

3. $H_0: \mu_d = 0$ vs $H_1: \mu_d \neq 0$

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} = 3.645422$$

$$W = \{|t| \geqslant t_{1-\alpha/2}(n-1)\}$$

$$t_{0.975}(7) = 2.365$$

接绝 H_0，两种方法有显著差异。

4. $H_0: \mu_d = 0$ vs $H_1: \mu_d \neq 0$

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} = 6.1555$$

$$p = P(|t| \geqslant t_{1-\alpha/2})$$

$$t_{0.975}(8) = 2.306$$

拒绝 H_0，两种测定方法有显著差异。

5. $\bar{d} \pm t_{1-\alpha/2}(n-1)\dfrac{s_d}{\sqrt{n}}$

即　$[-5.5054, 6.5054]$。

6. （1）$H_0: \mu_d \leqslant 0$ vs $H_1: \mu_d > 0$

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} = 5.47$$

$$W = \{t \leqslant t_{1-\alpha}(n-1)\}$$

$$t_{0.95}(14) = 1.761$$

接受 H_0，不能认为饮食和锻炼对降低胆固醇有显著作用。

（2）$p = P(t \geqslant t_0) = P(t \geqslant 5.47) \approx 0.00004$。

(3) $\bar{d}\pm t_{1-\alpha/2}(n-1)\frac{s_d}{\sqrt{n}}=[16.33, 37.41]$。

习题 4.5

1. $H_0:\sigma\leqslant 1 \quad \text{vs} \quad H_1:\sigma>1$

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}\sim\chi^2(n-1)$$
$$W=\{\chi^2\geqslant c\}$$
$$c=\chi^2_{1-\alpha}(n-1)$$
$$n=10,\ s^2=1.34$$
$$\chi_0^2=12.06<\chi^2_{0.95}(9)=16.92$$

因此接受原假设 H_0: $\sigma\leqslant 1$。

2. μ 的置信区间是 $\bar{x}\pm t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}}=2\pm 1.3934$，即为 $[0.6066, 3.3934]$。

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}}=1.7531$$
$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}}=3.9546$$

∴ σ 的置信区间是 $[1.7531, 3.9546]$

3. (1) μ 的置信下限为 $\bar{x}-t_{1-\alpha}(n-1)\frac{s}{\sqrt{n}}=40.69$；

(2) σ 的置信上限为 $\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}}=1.543$。

4. σ 的置信上限为 $\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}}=18.82$

5. 即 $\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}}=\frac{1}{4}\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\sqrt{\frac{(n-1)s^2}{16}}$

即 $\chi^2_{\alpha}(n-1)=16$

∴ $n\geqslant 10$

6. σ 的置信区间为 $\left[\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}}\right]$，代入 $s=0.37$，$\alpha=0.05$，可得置信区间为 $[0.31, 0.46]$，包含 0.35，因此接受 H_0。

7. (1) $H_0:\sigma\leqslant 0.02$ vs $H_1:\sigma>0.02$

$$\chi^2=\frac{(n-1)s^2}{\sigma^2}\sim\chi^2(n-1)$$
$$W=\{\chi^2\geqslant\chi^2_{1-\alpha}(n-1)\}$$
$$\chi_0^2=\frac{14\times 0.016^2}{0.02^2}=8.96$$

$\chi^2_{1-\alpha}(n-1)=\chi^2_{0.95}(14)=23.5$

接受原假设 H_0。

(2) $p=P(\chi^2\geqslant\chi_0^2)=0.8336$。

(3) σ 的置信下限是 $\sqrt{\dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}}=0.0123$，所以接受。

8. $\chi^2=\dfrac{Q}{\sigma^2}\sim\chi^2(n-5)$

∴ σ^2 的 0.95 置信区间为：

$$\left[\frac{Q}{\chi^2_{1-\alpha/2}(n-5)},\ \frac{Q}{\chi^2_{\alpha/2}(n-5)}\right]=[4.78,\ 15.58]$$

9. (1) $H_0:\sigma_1^2=\sigma_2^2$ vs $H_1:\sigma_1^2\neq\sigma_2^2$

$$F=\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}\sim F(n-1,\ m-1)$$

$$W=\{F\leqslant F_{\alpha/2}(n-1,m-1)\quad \text{或}\quad F\geqslant F_{1-\alpha/2}(n-1,m-1)\}$$

代入数据得

$$F_0=\frac{0.35}{0.4}=0.0875$$

$F_{0.025}(14,\ 16)=0.3421$，$F_{0.975}(14,\ 16)=2.8170$，所以有显著差异。

(2) $\dfrac{\sigma_1}{\sigma_2}$ 的置信区间为：

$$\left[\frac{s_1}{s_2}\frac{1}{\sqrt{F_{1-\alpha/2}(n-1,\ m-1)}},\ \frac{s_1}{s_2}\sqrt{F_{1-\alpha/2}(m-1,\ n-1)}\right]=[0.607,\ 1.46]$$

10. X_1，X_2，X_3，X_4 服从 $N(\mu,\ \sigma^2)$

∴ $\dfrac{X_1-X_2}{\sigma}\sim N(0,\ 1)$

$\dfrac{(X_1-X_2)^2}{\sigma^2}\sim\chi^2(1)$

同理$\dfrac{(X_3-X_4)^2}{\sigma^2}\sim\chi^2(1)$

∴ $\dfrac{(X_3-X_4)^2}{(X_1-X_2)}\sim F(1,\ 1)$

∴ $P\left(\dfrac{(X_3-X_4)^2}{(X_1-X_2)^2}<40\right)\approx 0.9$

11. (1) $\dfrac{\sigma_1}{\sigma_2}$的置信区间为：

$$\left[\frac{s_A}{s_B}\sqrt{\frac{1}{F_{1-\alpha/2}(n_1-1,\ n_2-1)}},\ \frac{s_A}{s_B}\sqrt{F_{1-\alpha/2}(n_2-1,\ n_1-1)}\right]$$

$=[0.72,\ 1.11]$

(2) $s_W^2=\dfrac{(n_1-1)s_A^2+\ (n_2-1)s_B^2}{n_1+n_2-2}$

$\mu_1-\mu_2$ 的置信区间为：

$$\bar{X}-\bar{Y}\pm t_{1-\alpha/2}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

即 $[-68.46, -11.54]$。

12. $H_0: \sigma_{女}\leqslant\sigma_{男}$ vs $H_1: \sigma_{女}>\sigma_{男}$

$$F=\frac{s_{女}^2/\sigma_{女}^2}{s_{男}^2/\sigma_{男}^2}\sim F(n_{女}-1, n_{男}-1)$$

$$W=\{F\geqslant F_{1-\alpha}(n_{女}-1, n_{男}-1)$$

代入数据

$$F=\frac{1.093^2}{0.914^2}=1.43$$

$$F_{1-\alpha}(n_{女}-1, n_{男}-1)=F_{0.99}(20, 24)=2.74$$

所以接受 H_0，男性散布程度不低于女性。(思考：试讨论 H_0：$\sigma_{男}\leqslant\sigma_{女}$，情况如何？如果使用检验统计量 $F_1=\frac{S_{男}^2/\sigma_{男}^2}{S_{女}^2/\sigma_{女}^2}$，情况如何？)

13. 甲$=1$，乙$=2$

$$H_0: \sigma_1\geqslant\sigma_2 \text{ vs } H_1: \sigma_1<\sigma_2$$

$$F=\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}\sim F(n_1-1, n_2-1)$$

$$W=\{F\leqslant F_\alpha(n_1-1, n_2-1)\}$$

代入数据

$$F_0=\frac{8.5^2}{6.75^2}=1.5857$$

$$F_\alpha(n_1-1, n_2-1)=F_{0.05}(50, 178)=0.6717$$

$$F_0>F_\alpha$$

因此接受 H_0，甲地价格波动比乙地大

习题 4.6

1. $H_0: p\geqslant 0.3$ vs $H_1: p<0.3$

$$P(T\geqslant 3)=1-P(T\leqslant 2)=1-\sum_{i=0}^{2}\binom{15}{i}0.3^i\times 0.7^{15-i}$$

$$=0.873>0.05$$

∴ 看法合适

2. $H_0: p\leqslant 0.01$ vs $H_1: p>0.01$

$$W=\{T\geqslant c'\}$$

即 $$1-\sum_{i=0}^{c'}\binom{n}{i}p_0^i(1-p_0)^{n-i}\leqslant\alpha$$

∴ $c'=2$

∴ $W=\{T\geqslant 2\}$

3. $H_0: p\geqslant 0.8$ vs $H_1: p<0.8$

$$u=\frac{T-np_0}{\sqrt{np_0(1-p_0)}}=\frac{\bar{x}-p_0}{\sqrt{p_0(1-p_0)/n}}\sim N(0,1)$$

$W=\{u\leqslant u_\alpha\}$

$u_{0.95}=1.64$

$u_{0.05}=-1.64$

$u_0=-5.94<u_\alpha$

∴拒绝 H_0，不支持研究者的观点

4. $H_0: p=0.1$ vs $H_1: p\neq 0.1$

$$u=\frac{\bar{x}-p_0}{\sqrt{p_0(1-p_0)/n}}\sim N(0,1)$$

$W=\{|u|\geqslant u_{1-\alpha/2}\}$

$u_{0.975}=1.96$

$u_0=1.12$

∴不拒绝 H_0，合格率仍为 10%

5. 次品率的置信区间为：

$$\bar{x}\pm u_{1-\alpha/2}\sqrt{p_0\ (1-p_0)\ /n}$$

即 [0.09，0.235]。

6. $H_0: p=0.7$ vs $H_1: p\neq 0.7$

p 的置信区间为：

$$\left[\hat{p}-u_{1-\alpha/2}\sqrt{\frac{\hat{p}\ (1-\hat{p})}{n}},\ \hat{p}+u_{1-\alpha/2}\sqrt{\frac{\hat{p}\ (1-\hat{p})}{n}}\right]=[0.620\,6,\ 0.679\,4]$$

因为 0.7 在置信区间外，故拒绝 H_0，不能认为喜欢酸味饮料的比率为 0.7。

7. 男性=1，女性=2

$H_0: p_1-p_2\geqslant 0$ vs $H_1: p_1-p_2<0$

$$u=\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}+\frac{\hat{p}(1-\hat{p})}{m}}}\sim N(0,1)$$

$$\hat{p}=\frac{n\ \hat{p}_1+m\ \hat{p}_2}{m+n}$$

$W=\{u\leqslant u_\alpha\}$

$u_{0.99}=2.326$

$u_{0.01}=-2.326$

$u_0=-2.233$

接受 H_0，女性色盲率不低于男性色盲率。

8. $H_0: p_1-p_2=0$ vs $H_1: p_1-p_2\neq 0$

$$u=\frac{\hat{p}_1-\hat{p}_2-(P_1-P_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)}}\dot{\sim}N(0,1)$$

$$\hat{p}=\frac{n\hat{p}_1+m\hat{p}_2}{m+n}$$

$$W=\{|u|\geqslant u_{1-\alpha/2}\}=\{|u|\geqslant u_{0.995}\}$$

$$u_{0.995}=2.576$$

$$u_0=-8.71$$

拒绝 H_0，两种肥料有显著差异。

9. $H_0: p_1-p_2=0$ vs $H_1: p_1-p_2\neq 0$

$$u=\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)}}\sim N(0,1)$$

$$\hat{p}=\frac{n\hat{p}_1+m\hat{p}_2}{m+n}$$

$$W=\{|u|>u_{1-\alpha/2}\}$$

$$u_0=2.01$$

$$u_{1-\alpha/2}=1.96$$

拒绝 H_0，制造方法对废品率有显著影响。

10. (1) $H_0: p_1-p_2=0$ vs $H_1: p_1-p_2\neq 0$

$$u_0=\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)}}=1.491$$

$$P=P\{|u|\geqslant u_0\}=0.136\,2$$

较大，不能拒绝 H_0，认为不合格品率相同合理。

(2) $p=0.136\,2$。

(3) 0.95 的置信区间为 $\hat{p}_1-\hat{p}_2\pm u_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})\frac{2}{n}}$，为 $[-0.007\,3, 0.054]$。

11. $$n=\left(\frac{u_{1-\alpha/2}\sqrt{\frac{(p_1+p_2)(q_1+q_2)}{2}}+u_{1-\beta}\sqrt{p_1q_1+p_2q_2}}{p_1-p_2}\right)^2$$

代入得 $n=787$。

12. 同第 11 题，可得 $n=973$，$n=673$。

13. p_A-p_B 的置信区间为：

$$\hat{p}_A-\hat{p}_B\pm u_{1-\alpha/2}\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n}+\frac{\hat{p}_B(1-\hat{p}_B)}{m}}$$

即 $[-0.163, 0.023]$。

习题 4.7

1. (1) $\Theta=\{(\mu_1, \mu_2, \sigma^2): -\infty<\mu_1, \mu_2<\infty, \sigma^2>0\}$

$\Theta_0=\{(\mu, \sigma^2): -\infty<\mu=\mu_1=\mu_2<\infty, \sigma^2>0\}$

(2) $$\max_{\Theta} L=\left\{\frac{n_1+n_2}{2\pi\left[\sum(x_{1i}-\bar{x}_1)^2+\sum(x_{2i}-\bar{x}_2)^2\right]}\right\}^{\frac{n_1+n_2}{2}} \mathrm{e}^{-\frac{n_1+n_2}{2}}$$

$$\max_{\Theta_0} L=\left\{\frac{n_1+n_2}{2\pi\left[\sum(x_{1i}-\bar{x}_1)^2+\sum(x_{1i}-\bar{x}_2)^2+\frac{n_1 n_2}{n_1+n_2}(\bar{x}_1-\bar{x}_2)^2\right]}\right\}^{\frac{n_1+n_2}{2}} \mathrm{e}^{-\frac{n_1+n_2}{2}}$$

(3) $$\lambda(\boldsymbol{x}_1, \boldsymbol{x}_2)=\left[1+\frac{\frac{n_1 n_2}{n_1+n_2}(\bar{x}_1-\bar{x}_2)^2}{\sum(x_{1i}-\bar{x}_1)^2+\sum(x_{2i}-\bar{x}_2)^2}\right]^{-\frac{n_1+n_2}{2}}, W=\{\lambda \leqslant c\}$$

(4) λ 是 $|t|$ 的严减函数。

2. $$\max_{\Theta_0} L=\left\{\frac{n}{2\pi\left[\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2+\sum\limits_j n_j(\bar{x}_j-\bar{x})^2\right]}\right\}^{\frac{n}{2}}$$

$$\max_{\Theta} L=\left\{\frac{n}{2\pi\left[\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2\right]}\right\}^{\frac{n}{2}}$$

$$\therefore \lambda(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k)=\left\{\frac{\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2+\sum\limits_j n_j(\bar{x}_j-\bar{\bar{x}})^2}{\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2}\right\}^{-\frac{n}{2}}, n=\sum_{j=1}^k n_j$$

$$=\left[1+\frac{k-1}{n-k} \frac{\sum\limits_j n_j(\bar{x}_j-\bar{\bar{x}})^2/(k-1)}{\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2/(n-k)}\right]^{-\frac{n}{2}}$$

$\bar{\bar{x}}$ 为 k 个样本的均值，用 $F(k-1, n-k)$ 获得水平为 α 的检验。

3. 当 $\sigma_1^2=\cdots=\sigma_k^2$ 时且未知时

$$\hat{\sigma}_j^2=\frac{1}{n_j} \sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2, j=1, \cdots k$$

$$\sigma_0^2=\frac{1}{\sum\limits_{j=1}^k n_j} \sum_{j=1}^k \hat{\sigma}_j^2$$

$$\therefore \max_{\Theta_0} L=\frac{1}{\left(2\pi\left(\sum\limits_{j=1}^k n_j \hat{\sigma}_j^2\right)/\sum\limits_{j=1}^k n_j\right)^{\frac{\sum n_j}{2}}} \mathrm{e}^{-\frac{\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2}{2\prod\limits_{j=1}^k(\sigma_j^2)}}$$

$$\max_{\Theta} L=\frac{1}{\prod\limits_{j=1}^k(2\pi \cdot \sigma_j^2)^{\frac{n_j}{2}}} \mathrm{e}^{-\frac{\sum\limits_j \sum\limits_i(x_{ji}-\bar{x}_j)^2}{2\prod\limits_{j=1}^k(\sigma_j^2)}}$$

$$\therefore \lambda(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k)=\frac{\prod_{j=1}^{k}(\sigma_j^2)^{\frac{n_j}{2}}}{\left(\sum_{j=1}^{k}n_j\hat{\sigma}_j^2\Big/\sum_{j=1}^{k}n_j\right)^{\sum n_j/2}}$$

当 σ_j 未知时 $\hat{\sigma}_j$ 为其估计，此时，由高等数理统计的结论知

$-2\ln\lambda \sim \chi^2(k-1)$

$\therefore W=\{-2\ln\lambda \geqslant \chi^2_{1-\alpha}(k-1)\}$

4. 见例4.7.4，样本来自正态分布。

习题 5.1

1.

i	$x_{(i)}$	$x_{(10+1-i)}$	a_i	$a_i\left[x_{(10+1-i)}-x_{(i)}\right]$
1	−1.2	3.7	0.573 9	2.812 11
2	−1	2.7	0.329 1	1.217 67
3	−0.6	2	0.214 1	0.556 66
4	−0.3	0.8	0.122 4	0.134 64
5	0	0.7	0.039 9	0.027 93

$$\sum_{i=1}^{5}a_i\left[x_{(10+1-i)}-x_{(i)}\right]=4.749\,01$$

$$Q=\sum_{i=1}^{10}(x_{(i)}-\bar{x})^2=24.376$$

$$\therefore W=\frac{(4.749\,01)^2}{24.376}=0.925$$

$n=10$，$W_{0.05}=0.842$，$W>W_{0.05}$

接受 H_0，即治疗前后血红蛋白的差服从正态分布。

2. 同第1题，在 $\alpha=0.05$ 下可认为抗压强度服从正态分布。

3. 同例5.1.2，在 $\alpha=0.05$ 下可认为该种人男子头颅的最大宽度服从正态分布。

4. (1) 同第1题：

$$W=\frac{\left[\sum_{i=1}^{12}a_i(x_{(26-i)}-x_{(i)})\right]^2}{\sum_{i=1}^{25}(x_{(i)}-\bar{x})^2}<W_{0.05}=0.918$$

拒绝 H_0，不服从正态分布

(2) 令 $Y_i=\lg(204-x_i)$

$$W=\frac{\left[\sum_{i=1}^{12}a_i(Y_{(26-i)}-Y_{(i)})\right]^2}{\sum_{i=1}^{25}(Y_{(i)}-\bar{Y})^2}>W_{0.05}=0.918$$

接受 H_0，认为服从正态分布。

习题 5.2

1. 来自 $N(0, 1)$。
2. 来自 exp(1/1 500)。
3. (1) $p=0.435$；
   (2) $p=0.0004$。
4. 不拒绝样本来自均匀分布 $U(0, 1\,440)$ 的假设。

习题 5.3

1. H_0：骰子是均匀的

$$\chi^2=\sum_{i=1}^{r}\frac{(Q_i-E_i)^2}{E_i}$$

式中,Q_i 为实际出现的频数；E_i 是在服从均匀分布的情况下的次数，$E_i=\frac{100}{6}$。

$$\begin{aligned}\therefore\ \chi^2 &=\frac{\left(13-\frac{100}{6}\right)^2}{100/6}+\frac{\left(14-\frac{100}{6}\right)^2}{100/6}+\frac{\left(20-\frac{100}{6}\right)^2}{100/6}+\frac{\left(17-\frac{100}{6}\right)^2}{100/6}+\frac{\left(15-\frac{100}{6}\right)^2}{100/6}\\&\quad+\frac{\left(21-\frac{100}{6}\right)^2}{100/6}\\&=3.2\end{aligned}$$

$\alpha=0.05$，$\chi^2_{1-\alpha}(5)=11.07$，$W=\{\chi^2\geqslant c\}$，$c=\chi^2_{1-\alpha}(r-1)$，$\chi^2<\chi^2_{1-\alpha}(5)$

接受原假设，即在 $\alpha=0.05$ 下，这颗骰子是均匀的。

2. 同第 1 题，在 $\alpha=0.05$ 下，认为 π 前 800 位数中 0～9 是等可能出现的。
3. 同第 1 题，认为职业病假人数在 5 个工作日上非均匀分布。
4. H_0：广告战前后各公司的市场占有率无显著变化

令 C 为其他公司

$$E_A=200\times\frac{45}{100}=90$$

$$E_B=200\times\frac{40}{100}=80$$

$$E_C=200\times\frac{15}{100}=30$$

$$\chi^2=\frac{(102-90)^2}{90}+\frac{(82-80)^2}{80}+\frac{(16-30)^2}{30}=8.18$$

$$W=\{\chi^2\geqslant\chi^2_{1-\alpha}(r-1)\}$$

$$\chi^2_{1-0.05}(2)=5.99$$

拒绝 H_0，广告战前后各公司的市场占有率有明显的变化。

5. 首先用观察数据 Q_i 获得泊松分布中未知参数 λ 的最大似然估计

$$\hat{\lambda}=\bar{x}=3.88$$

由此可计算出各泊松概率：

$$\hat{p}_j = \frac{\hat{\lambda}^j}{j!}e^{-\hat{\lambda}}, \ j = 0, 1, \cdots, 10$$

$$\hat{p}_{11} = \sum_{j=11}^{\infty} \frac{\hat{\lambda}}{j!}e^{-\hat{\lambda}}$$

由此计算出各期望频数 E_i，并计算统计量

$$\chi^2 = \sum_{i=0}^{11} \frac{(Q_i - E_i)^2}{E_i} < \chi^2_{0.9}(10) = 15.99$$

接受 H_0，认为服从泊松分布。

6. $\hat{\theta} = \frac{1}{162}\sum_{i=1}^{9} Q_i x_i$

$$\hat{p}_i = \frac{1}{\hat{\theta}}e^{-\frac{x_i}{\hat{\theta}}}$$

$$E_i = n\hat{p}_i$$

$$\chi^2 = \sum_{i=1}^{9} \frac{(Q_i - E_i)^2}{E_i} = 2.184 < \chi^2_{0.95}(8) = 15.51$$

∴ 不能拒绝 H_0，即 X 服从指数分布

7. $E_i = 200/10 = 20$

$$\chi^2 = \sum_{i=1}^{10} \frac{(Q_i - E_i)^2}{E_i} = 24.9 > \chi^2_{0.95}(9) = 16.92$$

拒绝 H_0，认为不具有随机性。

8. H_0：无影响

$$p_{ij} = p_i p_j$$

$$\hat{p}_{i\cdot} = \frac{Q_{i\cdot}}{\sum Q_{i\cdot}}$$

$$\hat{p}_{\cdot j} = \frac{Q_{\cdot j}}{\sum Q_{\cdot j}}$$

$$E_{ij} = np_{ij} = n\hat{p}_{i\cdot} \cdot \hat{P}_{\cdot j}$$

$$\chi^2 = \sum_{j=1}^{4}\sum_{i=1}^{4} \frac{(Q_{ij} - E_{ij})^2}{E_{ij}} = 14.395$$

$$\chi^2_{0.99}(9) = 21.67$$

不能拒绝 H_0

9. (1) 首先求 p 的最大似然估计，令 n_1=男正常，n_2=女正常，n_3=男色盲，n_4=女色盲

$$P(n_1, n_2, n_3, n_4) = \frac{n!}{n_1!\ n_2!\ n_3!\ n_4!} p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{n_4}$$

$$n=n_1+n_2+n_3+n_4$$

$$p_1=\frac{p}{2}$$

$$p_2=\frac{p^2}{2}+p(1-p)$$

$$p_3=\frac{1-p}{2}$$

$$p_4=\frac{(1-p)^2}{2}$$

即 $P(n_1, n_2, n_3, n_4)$

$$=\frac{n!}{n_1!\ n_2!\ n_3!\ n_4!}\left(\frac{p}{2}\right)^{n_1}\left[\frac{p^2}{2}+p(1-p)\right]^{n_2}\left(\frac{1-p}{2}\right)^{n_3}\left[\frac{(1-p)^2}{2}\right]^{n_4}$$

两边取对数

$$\begin{aligned}\ln P&=\ln\frac{n!}{n_1!\ n_2!\ n_3!\ n_4!}+n_1\ln\frac{p}{2}+n_2\ln\left[\frac{p^2}{2}+p(1-p)\right]\\&\quad+n_3\ln\frac{1-p}{2}+n_4\ln\frac{(1-p)^2}{2}\\&=\ln\frac{n!}{n_1!\ n_2!\ n_3!\ n_4!}+n_1\ln p-n_1\ln 2+n_2\ln(2p-p^2)-n_2\ln 2\\&\quad+n_3\ln(1-p)-n_3\ln 2+2n_4\ln(1-p)-n_4\ln 2\end{aligned}$$

两边对 p 对导

$$\frac{\partial\ln P}{\partial p}=\frac{n_1}{p}+n_2\frac{2-2p}{2p-p^2}-n_3\frac{1}{1-p}-2n_4\frac{1}{1-p}=0$$

即

$$(n_1+2n_2+n_3+2n_4)p^2-(3n_1+4n_2+2n_3+4n_4)p+2n_1+2n_2=0$$

$$n_1=442,\ n_2=514,\ n_3=38,\ n_4=6$$

代入可得

$$1\,520p^2-3\,482p+1\,912=0$$

$\hat{P}_{MLE}=0.913$，从而求得期望频数为：

	男	女
正常	456.5	496.2
色盲	43.5	3.8

从而

$$\chi^2=\frac{(456.5-442)^2}{456.5}+\frac{(496.2-514)^2}{496.2}+\frac{(43.5-38)^2}{43.5}+\frac{(3.8-6)^2}{3.8}=3.068$$

$\chi^2_{0.95}(2)=5.99$，$W=\{\chi^2\geqslant 5.99\}$

∴ 接受原假设，即数据和模型一致。

（2）H_0：性别与色盲独立

当 H_0 成立，求得其期望频数为：

	男	女
正常	458.9	497.1
色盲	21.1	22.9

$$\chi^2=\frac{(458.9-442)^2}{458.9}+\frac{(514-497.1)^2}{497.1}+\frac{(38-21.1)^2}{21.1}+\frac{(22.9-6)^2}{22.9}=27.21$$

$\chi^2_{0.95}(1)=3.84$

∴ 拒绝 H_0，不独立

10. (1) H_0：一致

$p_{1j}=p_{2j}=p_j,\quad j=1,2,3$

$$\hat{p}_j=\frac{O_{1j}+O_{2j}}{500+500}$$

∴ $\hat{p}_1=0.476$，$\hat{p}_2=0.281$，$\hat{p}_3=0.243$

∴ $E_{11}=238$，$E_{12}=140.5$，$E_{13}=121.5$，$E_{21}=238$，$E_{22}=140.5$，$E_{23}=121.5$

$$\therefore\ \chi^2=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(Q_{ij}-E_{ij})^2}{E_{ij}}=7.57$$

$\chi^2_{0.95}(2)=5.99$

拒绝原假设，两个三项分布并不相同。

11. 同第10题，在 $\alpha=0.01$ 下，四个二点分布间有显著差异。

12. H_0：性别与文化程度无关

H_0 成立时，其期望频数为：

	大专以上	中专技校	高中	初中及以下
男	36.82	128.87	651.714	1 023.6
女	23.18	81.13	410.286	644.404

$$\chi^2=\sum_{i=1}^{2}\sum_{j=1}^{4}\frac{(Q_{ij}-E_{ij})^2}{E_{ij}}=7.332$$

$\chi^2_{0.95}(6)=12.59$

$\chi^2<12.59$

即在 $\alpha=0.05$ 下，认为失业人员的性别与文化程度无关。

13. 同第12题，在 $\alpha=0.05$ 下，认为广告与人们对产品质量的评价无关。

14. 同第12题，在 $\alpha=0.05$ 下，认为该城市居民各年对社会热点问题看法有变化。

15. 同第12题，认为高血压与冠心病有相互影响。

16. 独立时其期望频数为：

	B	$\bar{B}$
A	$\frac{(a+b)(a+c)}{n}$	$\frac{(b+d)(a+b)}{n}$
$\bar{A}$	$\frac{(a+c)(c+d)}{n}$	$\frac{(b+d)(c+d)}{n}$

$$\chi^2=\frac{\left(\frac{(a+b)(a+c)}{n}-a\right)^2}{(a+b)(a+c)/n}+\frac{\left(\frac{(b+d)(a+b)}{n}-b\right)^2}{(b+d)(a+b)/n}$$
$$+\frac{\left(\frac{(a+c)(c+d)}{n}-c\right)^2}{(a+c)(c+d)/n}+\frac{\left(\frac{(b+d)(c+d)}{n}-d\right)^2}{(b+d)(c+d)/n}$$

展开即证。

17. 同第 12 题，在 $\alpha=0.05$ 下，认为废品率与制造方法有关。