# 习题 5.5

1．　设 $X_1, \cdots, X_n$ 是来自几何分布 $P\{X=x\}=\theta(1-\theta)^x, x=0,1,2,\cdots$ 的样本，证明 $T=\sum_{i=1}^{n} X_i$ 是充分统计量.

证：方法一：根据充分统计量的定义

样本联合概率函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\theta(1-\theta)^{x_i}=\theta^n(1-\theta)^{\sum_{i=1}^{n}x_i},$$

因 $X_i+1$ 的概率函数为 $P\{X_i+1=x\}=\theta(1-\theta)^x$，$x=1,2,\cdots$，即服从几何分布 $Ge(\theta)$，$i=1,2,\cdots,n$，

则根据几何分布与负二项分布的关系可知 $\sum_{i=1}^{n}(X_i+1)=T+n$ 服从负二项分布 $Nb(n,\theta)$，即概率函数为

$$P\{T+n=k\}=\binom{k-1}{n-1}\theta^n(1-\theta)^{k-n},\quad k=n,n+1,n+2,\cdots,$$

即 $T=\sum_{i=1}^{n}X_i$ 的概率函数为 $p_T(t;\theta)=\binom{t+n-1}{n-1}\theta^n(1-\theta)^t,\quad t=0,1,2,\cdots,$

可得在 $T=t$ 时，即 $t=\sum_{i=1}^{n}x_i$，$X_1,X_2,\cdots,X_n$ 的条件概率函数为

$$p(x_1,x_2,\cdots,x_n;\theta\,|\,T=t)=\frac{p(x_1,x_2,\cdots,x_n;\theta\}}{p_T(t;\theta)}=\frac{\theta^n(1-\theta)^{\sum_{i=1}^{n}x_i}}{\binom{t+n-1}{n-1}\theta^n(1-\theta)^t}=\frac{1}{\binom{t+n-1}{n-1}},$$

这与参数 $\theta$ 无关，

故根据充分统计量的定义可知 $T=\sum_{i=1}^{n}X_i$ 是 $\theta$ 的充分统计量.

方法二：根据因子分解定理

样本联合概率函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\theta(1-\theta)^{x_i}=\theta^n(1-\theta)^{\sum_{i=1}^{n}x_i},$$

因 $T=\sum_{i=1}^{n}X_i$，有 $t=\sum_{i=1}^{n}x_i$，即 $p(x_1,x_2,\cdots,x_n;\theta)=\theta^n(1-\theta)^t$，

取 $g(t;\theta)=\theta^n(1-\theta)^t$，$h(x_1,x_2,\cdots,x_n)=1$ 与参数 $\theta$ 无关，

故根据因子分解定理可知 $T=\sum_{i=1}^{n}X_i$ 是 $\theta$ 的充分统计量.

2．　设 $X_1, \cdots, X_n$ 是来自泊松分布 $P(\lambda)$ 的样本，证明 $T=\sum_{i=1}^{n}X_i$ 是充分统计量.

证：方法一：根据充分统计量的定义

样本联合概率函数

$$p(x_1,x_2,\cdots,x_n;\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{x_1!x_2!\cdots x_n!}\mathrm{e}^{-n\lambda}=\lambda^{\sum_{i=1}^{n}x_i}\mathrm{e}^{-n\lambda}\cdot\frac{1}{x_1!x_2!\cdots x_n!},$$

根据泊松分布的可加性可知 $T=\sum_{i=1}^{n}X_i$ 服从泊松分布 $P(n\lambda)$，即概率函数为

$$p_T(t;\lambda)=\frac{(n\lambda)^t}{t!}\mathrm{e}^{-n\lambda},\quad t=0,1,2,\cdots,$$

可得在 $T=t$ 时，即 $t=\sum_{i=1}^{n}x_i$，$X_1,X_2,\cdots,X_n$ 的条件概率函数为

$$p(x_1,x_2,\cdots,x_n;\theta\,|\,T=t)=\frac{p(x_1,x_2,\cdots,x_n;\theta\}}{p_T(t;\theta)}=\frac{\lambda^{\sum_{i=1}^{n}x_i}\mathrm{e}^{-n\lambda}\cdot\dfrac{1}{x_1!x_2!\cdots x_n!}}{\dfrac{n^t\lambda^t}{t!}\mathrm{e}^{-n\lambda}}=\frac{t!}{n^t\cdot x_1!x_2!\cdots x_n!},$$

这与参数$\lambda$ 无关，

故根据充分统计量的定义可知 $T=\sum_{i=1}^{n}X_i$ 是$\lambda$ 的充分统计量.

方法二：根据因子分解定理

样本联合概率函数

$$p(x_1,x_2,\cdots,x_n;\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}\mathrm{e}^{-\lambda}=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{x_1!x_2!\cdots x_n!}\mathrm{e}^{-n\lambda}=\lambda^{\sum_{i=1}^{n}x_i}\mathrm{e}^{-n\lambda}\cdot\frac{1}{x_1!x_2!\cdots x_n!},$$

因 $T=\sum_{i=1}^{n}X_i$，有 $t=\sum_{i=1}^{n}x_i$，即 $p(x_1,x_2,\cdots,x_n;\lambda)=\lambda^t\mathrm{e}^{-n\lambda}\cdot\dfrac{1}{x_1!x_2!\cdots x_n!}$，

取 $g(t;\lambda)=\lambda^t\mathrm{e}^{-n\lambda}$，$h(x_1,x_2,\cdots,x_n)=\dfrac{1}{x_1!x_2!\cdots x_n!}$ 与参数$\lambda$ 无关，

故根据因子分解定理可知 $T=\sum_{i=1}^{n}X_i$ 是$\lambda$ 的充分统计量.

3． 设总体为如下离散型分布，

| $X$ | $a_1$ | $a_2$ | $\cdots$ | $a_k$ |
|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | $\cdots$ | $p_k$ |

$X_1,\cdots,X_n$ 是来自该总体的样本，

（1）证明次序统计量 $(X_{(1)},\cdots,X_{(n)})$ 是充分统计量.

（2）以 $n_j$ 表示 $X_1,\cdots,X_n$ 中等于 $a_j$ 的个数，证明 $(n_1,\cdots,n_k)$ 是充分统计量.

证：设样本 $(X_1,X_2,\cdots,X_n)$ 中有 $n_1$ 个 $a_1$，$n_2$ 个 $a_2$，$\cdots$，$n_k$ 个 $a_k$，

显然次序统计量 $(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 中同样有 $n_1$ 个 $a_1$ ，$n_2$ 个 $a_2$ ，…，$n_k$ 个 $a_k$ ，
样本联合概率函数

$$p(x_1, x_2, \cdots, x_n; p_1, p_2, \cdots, p_k) = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k} ,$$

（2）因 $T_2 = (n_1, \cdots, n_k)$，取 $g(n_1, n_2, \cdots, n_k; p_1, p_2, \cdots, p_k) = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$ ，$h(x_1, x_2, \cdots, x_n) = 1$，

故根据因子分解定理可知 $T_2 = (n_1, n_2, \cdots, n_k)$ 是 $(p_1, p_2, \cdots, p_k)$ 的充分统计量；

（1）因 $T_1 = (X_{(1)}, X_{(2)}, \cdots, X_{(n)})$，显然$(n_1, n_2, \cdots, n_k)$ 与 $(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 一一对应，
故由第（2）小题结论知 $T_1 = (X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 是 $(p_1, p_2, \cdots, p_k)$ 的充分统计量.

4． 设 $X_1, \cdots, X_n$ 是来自正态分布 $N(\mu, 1)$ 的样本，证明 $T = \sum_{i=1}^{n} X_i$ 是充分统计量

证：方法一：根据充分统计量的定义
样本联合密度函数

$$p(x_1, x_2, \cdots, x_n; \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}} = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i^2 - 2\mu x_i + \mu^2)} = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \mu \sum_{i=1}^{n} x_i - \frac{1}{2} n\mu^2} ,$$

根据正态分布的可加性可知 $T = \sum_{i=1}^{n} X_i$ 服从正态分布 $N(n\mu, n)$，即密度函数为

$$p_T(t) = \frac{1}{\sqrt{2\pi} \cdot \sqrt{n}} e^{-\frac{(t-n\mu)^2}{2n}} = \frac{1}{\sqrt{2\pi} \cdot \sqrt{n}} e^{-\frac{t^2}{2n} + \mu t - \frac{1}{2} n\mu^2} ,$$

可得在 $T = t$ 时，即 $t = \sum_{i=1}^{n} x_i$ ，$X_1, X_2, \cdots, X_n$ 的条件概率函数为

$$p(x_1, x_2, \cdots, x_n; \mu \mid T = t) = \frac{p(x_1, x_2, \cdots, x_n; \mu)}{p_T(t)}$$

$$= \frac{\frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \mu \sum_{i=1}^{n} x_i - \frac{1}{2} n\mu^2}}{\frac{1}{\sqrt{2\pi} \cdot \sqrt{n}} e^{-\frac{t^2}{2n} + \mu t - \frac{1}{2} n\mu^2}} = \frac{\sqrt{n}}{(\sqrt{2\pi})^{n-1}} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \frac{t^2}{2n}} = \frac{\sqrt{n}}{(\sqrt{2\pi})^{n-1}} e^{-\frac{1}{2}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)} ,$$

这与参数 $\mu$ 无关，

故根据充分统计量的定义可知 $T = \sum_{i=1}^{n} X_i$ 是 $\mu$ 的充分统计量.

方法二：根据因子分解定理
样本联合密度函数

$$p(x_1, x_2, \cdots, x_n; \mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}} = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i^2 - 2\mu x_i + \mu^2)} = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \mu \sum_{i=1}^{n} x_i - \frac{1}{2} n\mu^2} ,$$

因 $T = \sum_{i=1}^{n} X_i$ ，有 $t = \sum_{i=1}^{n} x_i$ ，即 $p(x_1, x_2, \cdots, x_n; \mu) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \mu t - \frac{1}{2} n\mu^2} = \frac{1}{(\sqrt{2\pi})^n} e^{\mu t - \frac{1}{2} n\mu^2} \cdot e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2}$ ,

取 $g(t;\mu)=\dfrac{1}{(\sqrt{2\pi})^n}\mathrm{e}^{\mu t-\frac{1}{2}n\mu^2}$， $h(x_1,x_2,\cdots,x_n)=\mathrm{e}^{-\frac{1}{2}\sum\limits_{i=1}^{n}x_i^2}$ 与参数$\mu$ 无关，

故根据因子分解定理可知$T=\sum\limits_{i=1}^{n}X_i$ 是$\mu$ 的充分统计量．

5．设 $X_1$，…, $X_n$ 是来自 $p(x;\theta)=\theta x^{\theta-1}$，$0<x<1$，$\theta>0$ 的样本，试给出一个充分统计量．

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}\mathrm{I}_{0<x_i<1}=\theta^n(x_1x_2\cdots x_n)^{\theta-1}\mathrm{I}_{0<x_1,x_2,\cdots,x_n<1},$$

令 $T=X_1X_2\cdots X_n$，有 $t=x_1x_2\cdots x_n$，即 $p(x_1,x_2,\cdots,x_n;\theta)=\theta^n t^{\theta-1}\mathrm{I}_{0<x_1,x_2,\cdots,x_n<1}$，

取 $g(t;\theta)=\theta^n t^{\theta-1}$，$h(x_1,x_2,\cdots,x_n)=\mathrm{I}_{0<x_1,x_2,\cdots,x_n<1}$ 与参数$\theta$ 无关，

故根据因子分解定理可知 $T=X_1X_2\cdots X_n$ 是$\theta$ 的充分统计量．

6．设 $X_1$，…, $X_n$ 是来自韦布尔分布 $p(x;\theta)=mx^{m-1}\theta^{-m}\mathrm{e}^{-(x/\theta)^m}$，$x>0$，$\theta>0$ 的样本（$m>0$ 已知），试给出一个充分统计量．

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}mx_i^{m-1}\theta^{-m}\mathrm{e}^{-(x_i/\theta)^m}\mathrm{I}_{x_i>0}=m^n(x_1x_2\cdots x_n)^{m-1}\theta^{-mn}\mathrm{e}^{-\sum\limits_{i=1}^{n}(x_i/\theta)^m}\mathrm{I}_{x_1,x_2,\cdots,x_n>0}$$

$$=\theta^{-mn}\mathrm{e}^{-\frac{1}{\theta^m}\sum\limits_{i=1}^{n}x_i^m}\cdot m^n(x_1x_2\cdots x_n)^{m-1}\mathrm{I}_{x_1,x_2,\cdots,x_n>0},$$

令 $T=\sum\limits_{i=1}^{n}X_i^m$，有 $t=\sum\limits_{i=1}^{n}x_i^m$，即 $p(x_1,x_2,\cdots,x_n;\theta)=\theta^{-mn}\mathrm{e}^{-\frac{1}{\theta^m}t}\cdot m^n(x_1x_2\cdots x_n)^{m-1}\mathrm{I}_{x_1,x_2,\cdots,x_n>0}$，

取 $g(t;\theta)=\theta^{-mn}\mathrm{e}^{-\frac{1}{\theta^m}t}$，$h(x_1,x_2,\cdots,x_n)=m^n(x_1x_2\cdots x_n)^{m-1}\mathrm{I}_{x_1,x_2,\cdots,x_n>0}$ 与参数$\theta$ 无关，

故根据因子分解定理知$T=\sum\limits_{i=1}^{n}X_i^m$ 是$\theta$ 的充分统计量．

7．设 $X_1$，…, $X_n$ 是来 Pareto 分布 $p(x;\theta)=\theta a^\theta x^{-(\theta+1)}$，$x>a$，$\theta>0$ 的样本（$a>0$ 已知），试给出一个充分统计量．

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\theta a^\theta x_i^{-(\theta+1)}\mathrm{I}_{x_i>a}=\theta^n a^{n\theta}(x_1x_2\cdots x_n)^{-(\theta+1)}\mathrm{I}_{x_1,x_2,\cdots,x_n>a},$$

令 $T=X_1X_2\cdots X_n$，有 $t=x_1x_2\cdots x_n$，即 $p(x_1,x_2,\cdots,x_n;\theta)=\theta^n a^{n\theta}t^{-(\theta+1)}\mathrm{I}_{x_1,x_2,\cdots,x_n>a}$，

取 $g(t;\theta)=\theta^n a^{n\theta}t^{-(\theta+1)}$，$h(x_1,x_2,\cdots,x_n)=\mathrm{I}_{x_1,x_2,\cdots,x_n>a}$ 与参数$\theta$ 无关，

故根据因子分解定理知 $T=X_1X_2\cdots X_n$ 是$\theta$ 的充分统计量．

8．设 $X_1, \cdots, X_n$ 是来自 Laplace 分布 $p(x;\theta)=\dfrac{1}{2\theta}e^{-|x|/\theta}$，$\theta>0$ 的样本，试给出一个充分统计量.

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\mu)=\prod_{i=1}^{n}\frac{1}{2\theta}\mathrm{e}^{-\frac{|x_i|}{\theta}}=\frac{1}{(2\theta)^n}\mathrm{e}^{-\frac{1}{\theta}\sum_{i=1}^{n}|x_i|},$$

令 $T=\sum_{i=1}^{n}|X_i|$，有 $t=\sum_{i=1}^{n}|x_i|$，即 $p(x_1,x_2,\cdots,x_n;\mu)=\dfrac{1}{(2\theta)^n}\mathrm{e}^{-\frac{1}{\theta}t}$，

取 $g(t;\theta)=\dfrac{1}{(2\theta)^n}\mathrm{e}^{-\frac{1}{\theta}t}$，$h(x_1,x_2,\cdots,x_n)=1$ 与参数 $\theta$ 无关，

故根据因子分解定理知 $T=\sum_{i=1}^{n}|X_i|$ 是 $\theta$ 的充分统计量.

9．设 $X_1, \cdots, X_n$ 独立同分布，$X_1$ 服从以下分布，求相应的充分统计量：

（1）负二项分布 $X_1\sim p(x_1;\theta)=\dbinom{x_1+r-1}{r-1}\theta^r(1-\theta)^{x_1}$，$x_1=0,1,2,\cdots$，$r$ 已知；

（2）离散均匀分布 $X_1\sim p(x_1;m)=\dfrac{1}{m}$，$x_1=1,2,\cdots,m$，$m$ 未知；

（3）对数正态分布 $X_1\sim p(x_1;\mu,\sigma)=\dfrac{1}{\sqrt{2\pi}\sigma x_1}\exp\left\{-\dfrac{1}{2\sigma^2}(\ln x_1-\mu)^2\right\}$，$x_1>0$；

（4）瑞利（Rayleigh）分布 $X_1\sim p(x_1;\mu,\sigma)=2\lambda x_1\,\mathrm{e}^{-\lambda x_1^2}\cdot \mathrm{I}_{x_1\ge 0}$.

注：第（4）小题有误，密度函数应为 $p(x_1;\lambda)$，即参数应为 $\lambda$，而不是 $\mu,\sigma$.

解：（1）样本联合密度函数为

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\binom{x_i+r-1}{r-1}\theta^r(1-\theta)^{x_i}=\theta^{nr}(1-\theta)^{\sum_{i=1}^{n}x_i}\cdot\prod_{i=1}^{n}\binom{x_i+r-1}{r-1},$$

令 $T=\sum_{i=1}^{n}X_i$，有 $t=\sum_{i=1}^{n}x_i$，即 $p(x_1,x_2,\cdots,x_n;\theta)=\theta^{nr}(1-\theta)^t\cdot\prod_{i=1}^{n}\dbinom{x_i+r-1}{r-1}$，

取 $g(t;\theta)=\theta^{nr}(1-\theta)^t$，$h(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}\dbinom{x_i+r-1}{r-1}$ 与参数 $\theta$ 无关，

故根据因子分解定理知 $T=\sum_{i=1}^{n}X_i$ 是参数 $\theta$ 的充分统计量；

（2）样本联合密度函数为

$$p(x_1,x_2,\cdots,x_n;m)=\prod_{i=1}^{n}\frac{1}{m}\cdot\mathrm{I}_{1\le x_i\le m,\,x_i\text{为整数}}=\frac{1}{m^n}\cdot\mathrm{I}_{1\le x_1,x_2,\cdots,x_n\le m,\,x_1,x_2,\cdots,x_n\text{为整数}},$$

$$=\frac{1}{m^n}\cdot\mathrm{I}_{1\le x_{(1)}\le x_{(n)}\le m,\,x_1,x_2,\cdots,x_n\text{为整数}}=\frac{1}{m^n}\cdot\mathrm{I}_{x_{(n)}\le m}\cdot\mathrm{I}_{x_{(1)}\ge 1,\,x_1,x_2,\cdots,x_n\text{为整数}},$$

令 $T = X_{(n)} = \max\limits_{1\le i\le n}\{X_i\}$，有 $t = x_{(n)}$，即 $p(x_1, x_2, \cdots, x_n; m) = \dfrac{1}{m^n}\cdot \mathrm{I}_{t\le m}\cdot \mathrm{I}_{x_{(1)}\ge 1,\, x_1, x_2, \cdots, x_n\text{为整数}}$，

取 $g(t; m) = \dfrac{1}{m^n}\cdot \mathrm{I}_{t\le m}$，$h(x_1, x_2, \cdots, x_n) = \mathrm{I}_{x_{(1)}\ge 1,\, x_1, x_2, \cdots, x_n\text{为整数}}$ 与参数 $m$ 无关，

故根据因子分解定理知 $T = X_{(n)}$是参数 $m$ 的充分统计量；

（3）样本联合密度函数为

$$p(x_1, x_2, \cdots, x_n; \mu, \sigma) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma x_i}\exp\left\{-\frac{1}{2\sigma^2}(\ln x_i - \mu)^2\right\}\cdot \mathrm{I}_{x_i>0}$$

$$= \frac{1}{(\sqrt{2\pi}\sigma)^n x_1 x_2\cdots x_n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln^2 x_i - 2\mu\ln x_i + \mu^2)\right\}\cdot \mathrm{I}_{x_1, x_2, \cdots, x_n>0}$$

$$= \frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}\ln^2 x_i - 2\mu\sum_{i=1}^{n}\ln x_i + n\mu^2\right)\right\}\cdot\frac{1}{x_1 x_2\cdots x_n}\mathrm{I}_{x_1, x_2, \cdots, x_n>0},$$

令 $T_1 = \sum\limits_{i=1}^{n}\ln X_i$，$T_2 = \sum\limits_{i=1}^{n}\ln^2 X_i$，有 $t_1 = \sum\limits_{i=1}^{n}\ln x_i$，$t_2 = \sum\limits_{i=1}^{n}\ln^2 x_i$，

则 $p(x_1, x_2, \cdots, x_n; \mu, \sigma) = \dfrac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\dfrac{1}{2\sigma^2}(t_2 - 2\mu t_1 + n\mu^2)\right\}\cdot\dfrac{1}{x_1 x_2\cdots x_n}\cdot \mathrm{I}_{x_1, x_2, \cdots, x_n>0}$，

取 $g(t; \mu, \sigma) = \dfrac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\dfrac{1}{2\sigma^2}(t_2 - 2\mu t_1 + n\mu^2)\right\}$，

$h(x_1, x_2, \cdots, x_n) = \dfrac{1}{x_1 x_2\cdots x_n}\cdot \mathrm{I}_{x_1, x_2, \cdots, x_n>0}$ 与参数$\mu, \sigma$ 无关，

故根据因子分解定理知 $(T_1, T_2) = \left(\sum\limits_{i=1}^{n}\ln X_i, \sum\limits_{i=1}^{n}\ln^2 X_i\right)$是参数$(\mu, \sigma)$的充分统计量；

（4）样本联合密度函数为

$$p(x_1, x_2, \cdots, x_n; \lambda) = \prod_{i=1}^{n}2\lambda x_i\,\mathrm{e}^{-\lambda x_i^2}\cdot \mathrm{I}_{x_i>0} = 2^n\lambda^n x_1 x_2\cdots x_n\,\mathrm{e}^{-\lambda\sum_{i=1}^{n}x_i^2}\cdot \mathrm{I}_{x_1, x_2, \cdots, x_n>0},$$

令 $T = \sum\limits_{i=1}^{n}X_i^2$，有 $t = \sum\limits_{i=1}^{n}x_i^2$，即 $p(x_1, x_2, \cdots, x_n; \lambda) = 2^n\lambda^n\,\mathrm{e}^{-\lambda t}\cdot x_1 x_2\cdots x_n \mathrm{I}_{x_1, x_2, \cdots, x_n>0}$，

取 $g(t; \lambda) = 2^n\lambda^n\,\mathrm{e}^{-\lambda t}$，$h(x_1, x_2, \cdots, x_n) = x_1 x_2\cdots x_n\cdot \mathrm{I}_{x_1, x_2, \cdots, x_n>0}$ 与参数$\lambda$无关，

故根据因子分解定理知 $T = \sum\limits_{i=1}^{n}X_i^2$ 是参数$\lambda$的充分统计量.

10．设 $X_1, \cdots, X_n$ 是来自正态分布 $N(\mu, \sigma^2)$的样本.

（1）在$\mu$ 已知时给出$\sigma^2$ 的一个充分统计量；

（2）在$\sigma^2$ 已知时给出$\mu$ 的一个充分统计量.

解：因总体密度函数为

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

则样本联合密度函数为

$$p(x_1,x_2,\cdots,x_n;\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2},$$

（1）在$\mu$ 已知时，令$T_1=\sum_{i=1}^{n}(X_i-\mu)^2$，有$t=\sum_{i=1}^{n}(x_i-\mu)^2$，即$p(x_1,x_2,\cdots,x_n;\sigma^2)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{t}{2\sigma^2}}$，

取$g(t;\sigma^2)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{t}{2\sigma^2}}$，$h(x_1,x_2,\cdots,x_n)=1$ 与参数$\sigma^2$无关，

故根据因子分解定理知$T_1=\sum_{i=1}^{n}(X_i-\mu)^2$ 是参数$\sigma^2$的充分统计量；

（2）在$\sigma^2$ 已知时，

$$p(x_1,x_2,\cdots,x_n;\mu)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)}=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}x_i^2-2\mu\sum_{i=1}^{n}x_i+n\mu^2\right)}$$

$$=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)}=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i}\cdot\mathrm{e}^{-\frac{n\mu^2}{2\sigma^2}}\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2},$$

令$T_2=\sum_{i=1}^{n}X_i$，有$t=\sum_{i=1}^{n}x_i$，即$p(x_1,x_2,\cdots,x_n;\mu)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{\frac{\mu}{\sigma^2}t}\cdot\mathrm{e}^{-\frac{n\mu^2}{2\sigma^2}}\cdot\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2}$，

取$g(t;\mu)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\mathrm{e}^{\frac{\mu}{\sigma^2}t}\cdot\mathrm{e}^{-\frac{n\mu^2}{2\sigma^2}}$，$h(x_1,x_2,\cdots,x_n)=\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2}$ 与参数$\mu$ 无关，

故根据因子分解定理知$T_2=\sum_{i=1}^{n}X_i$ 是参数$\mu$ 的充分统计量.

11．设$X_1,\cdots,X_n$是来自均匀分布$U(\theta_1,\theta_2)$的样本，试给出一个充分统计量.

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\frac{1}{\theta_2-\theta_1}I_{\theta_1<x_i<\theta_2}=\frac{1}{(\theta_2-\theta_1)^n}I_{\theta_1<x_1,x_2,\cdots,x_n<\theta_2}=\frac{1}{(\theta_2-\theta_1)^n}I_{\theta_1<x_{(1)}\le x_{(n)}<\theta_2},$$

令$(T_1,T_2)=(X_{(1)},X_{(n)})$，有$(t_1,t_2)=(x_{(1)},x_{(n)})$，即$p(x_1,x_2,\cdots,x_n;\theta)=\frac{1}{(\theta_2-\theta_1)^n}I_{\theta_1<t_1\le t_2<\theta_2}$，

取$g(t_1,t_2;\theta_1,\theta_2)=\frac{1}{(\theta_2-\theta_1)^n}I_{\theta_1<t_1\le t_2<\theta_2}$，$h(x_1,x_2,\cdots,x_n)=1$ 与参数$\theta_1,\theta_2$无关，

故根据因子分解定理知$(T_1,T_2)=(X_{(1)},X_{(n)})$是$(\theta_1,\theta_2)$的充分统计量.

12．设$X_1,\cdots,X_n$是来自均匀分布$U(\theta,2\theta)$，$\theta>0$的样本，试给出充分统计量.

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}I_{\theta<x_i<2\theta}=\frac{1}{\theta^n}I_{\theta<x_1,x_2,\cdots,x_n<2\theta}=\frac{1}{\theta^n}I_{\theta<x_{(1)}\le x_{(n)}<2\theta},$$

令 $(T_1, T_2)=(X_{(1)}, X_{(n)})$，有$(t_1, t_2)=(x_{(1)}, x_{(n)})$，即 $p(x_1, x_2, \cdots, x_n;\theta)=\dfrac{1}{\theta^n} I_{\theta<t_1\le t_2<2\theta}$

取 $g(t_1, t_2;\theta)=\dfrac{1}{\theta^n} I_{\theta<t_1\le t_2<2\theta}$，$h(x_1, x_2, \cdots, x_n)=1$ 与参数$\theta$无关，

故根据因子分解定理知$(T_1, T_2)=(X_{(1)}, X_{(n)})$是$\theta$ 的充分统计量.

13．设 $X_1, \cdots, X_n$ 来自伽玛分布族$\{Ga(\alpha, \lambda)\mid \alpha>0, \lambda>0\}$的一个样本，寻求$(\alpha, \lambda)$的充分统计量.

解：总体 $X$ 的密度函数为

$$p(x;\alpha, \lambda)=\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\lambda x} \mathrm{I}_{x>0},$$

样本联合密度函数为

$$p(x_1, x_2, \cdots, x_n;\alpha, \lambda)=\prod_{i=1}^n \frac{\lambda^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} \mathrm{e}^{-\lambda x_i} \mathrm{I}_{x_i>0}=\frac{\lambda^{n\alpha}}{\Gamma(\alpha)^n}(x_1x_2\cdots x_n)^{\alpha-1} \mathrm{e}^{-\lambda\sum\limits_{i=1}^n x_i} \mathrm{I}_{x_1, x_2, \cdots, x_n>0},$$

令 $(T_1, T_2)=\left(X_1X_2\cdots X_n, \sum\limits_{i=1}^n X_i\right)$，有 $(t_1, t_2)=\left(x_1x_2\cdots x_n, \sum\limits_{i=1}^n x_i\right)$，

则 $p(x_1, x_2, \cdots, x_n;\alpha, \lambda)=\dfrac{\lambda^{n\alpha}}{\Gamma(\alpha)^n} t_1^{\alpha-1} \mathrm{e}^{-\lambda t_2} \mathrm{I}_{x_1, x_2, \cdots, x_n>0}$，

取 $g(t_1, t_2;\alpha, \lambda)=\dfrac{\lambda^{n\alpha}}{\Gamma(\alpha)^n} t_1^{\alpha-1} \mathrm{e}^{-\lambda t_2}$，$h(x_1, x_2, \cdots, x_n)=\mathrm{I}_{x_1, x_2, \cdots, x_n>0}$ 与参数$\alpha, \lambda$无关，

故 $(T_1, T_2)=\left(X_1X_2\cdots X_n, \sum\limits_{i=1}^n X_i\right)$是参数$(\alpha, \lambda)$的充分统计量.

14．设 $X_1, \cdots, X_n$ 是来自贝塔分布族$\{Be(a, b)\mid a>0, b>0\}$的一个样本，寻求$(a, b)$的充分统计量.

解：总体 $X$ 的密度函数为

$$p(x;a, b)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \mathrm{I}_{0<x<1},$$

样本联合密度函数

$$p(x_1, x_2, \cdots, x_n;a, b)=\prod_{i=1}^n p(x_i;a, b)=\prod_{i=1}^n \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x_i^{a-1}(1-x_i)^{b-1} \mathrm{I}_{0<x_i<1}$$

$$=\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]^n \left(\prod_{i=1}^n x_i\right)^{a-1} \left[\prod_{i=1}^n (1-x_i)\right]^{b-1} \mathrm{I}_{0<x_1, x_2, \cdots, x_n<1},$$

令 $(T_1, T_2)=\left(\prod\limits_{i=1}^n X_i, \prod\limits_{i=1}^n (1-X_i)\right)$，有 $(t_1, t_2)=\left(\prod\limits_{i=1}^n x_i, \prod\limits_{i=1}^n (1-x_i)\right)$，

则 $p(x_1, x_2, \cdots, x_n;a, b)=\left[\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]^n t_1^{a-1} t_2^{b-1}\cdot \mathrm{I}_{0<x_1, x_2, \cdots, x_n<1}$，

取 $g(t_1, t_2;a, b)=\left[\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right]^n t_1^{a-1} t_2^{b-1}$，$h(x_1, x_2, \cdots, x_n)=\mathrm{I}_{0<x_1, x_2, \cdots, x_n<1}$ 与参数 $a, b$ 无关，

故根据因子分解定理知 $(T_1,T_2)=\left(\prod_{i=1}^{n}X_i,\prod_{i=1}^{n}(1-X_i)\right)$ 是 $a, b$ 的充分统计量.

15．若 $X=(X_1,\cdots,X_n)$ 为从分布族 $f(x;\theta)=C(\theta)\exp\left\{\sum_{i=1}^{k}Q_i(\theta)T_i(x)\right\}h(x)$ 中抽取的简单样本，试证

$$T(X)=\left(\sum_{j=1}^{n}T_1(X_j),\cdots,\sum_{j=1}^{n}T_k(X_j)\right)$$

为充分统计量.

证：样本联合密度函数为

$$p(x_1,x_2,\cdots,x_n;\theta)=\prod_{j=1}^{n}C(\theta)\exp\left\{\sum_{i=1}^{k}Q_i(\theta)T_i(x_j)\right\}h(x_j)$$

$$=C(\theta)^n\exp\left\{\sum_{j=1}^{n}\sum_{i=1}^{k}Q_i(\theta)T_i(x_j)\right\}\cdot\prod_{j=1}^{n}h(x_j)=C(\theta)^n\exp\left\{\sum_{i=1}^{k}Q_i(\theta)\sum_{j=1}^{n}T_i(x_j)\right\}\cdot\prod_{j=1}^{n}h(x_j),$$

因 $T(X)=\left(\sum_{j=1}^{n}T_1(X_j),\cdots,\sum_{j=1}^{n}T_k(X_j)\right)$，有 $T(x)=(t_1,\cdots,t_k)=\left(\sum_{j=1}^{n}T_1(x_j),\cdots,\sum_{j=1}^{n}T_k(x_j)\right)$,

则 $p(x_1,x_2,\cdots,x_n;\theta)=C(\theta)^n\exp\left\{\sum_{i=1}^{k}Q_i(\theta)t_i\right\}\cdot\prod_{j=1}^{n}h(x_j)$,

取 $g(T(x);\theta)=C(\theta)^n\exp\left\{\sum_{i=1}^{k}Q_i(\theta)t_i\right\}$，$h(x_1,x_2,\cdots,x_n)=\prod_{j=1}^{n}h(x_j)$ 与参数 $\theta$ 无关,

故 $T(X)=\left(\sum_{j=1}^{n}T_1(X_j),\cdots,\sum_{j=1}^{n}T_k(X_j)\right)$ 为参数 $\theta$ 的充分统计量.

16．设 $X_1,\cdots,X_n$ 是来自正态总体 $N(\mu,\sigma_1^2)$ 的样本，$Y_1,\cdots,Y_m$ 是来自另一正态总体 $N(\mu,\sigma_2^2)$ 的样本，这两个样本相互独立，试给出 $(\mu,\sigma_1^2,\sigma_2^2)$ 的充分统计量.

解：两个总体的密度函数分别为

$$p_X(x;\mu,\sigma_1^2)=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma_1^2}},\quad p_Y(y;\mu,\sigma_2^2)=\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu)^2}{2\sigma_2^2}},$$

全部样本的联合密度函数为

$$p(x_1,\cdots,x_n,y_1,\cdots,y_m;\mu,\sigma_1^2,\sigma_2^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x_i-\mu)^2}{2\sigma_1^2}}\cdot\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y_j-\mu)^2}{2\sigma_2^2}}$$

$$=\frac{1}{(\sqrt{2\pi})^{n+m}\sigma_1^n\sigma_2^m}\mathrm{e}^{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)-\frac{1}{2\sigma_2^2}\sum_{j=1}^{m}(y_j^2-2\mu y_j+\mu^2)}$$

$$= \frac{1}{(\sqrt{2\pi})^{n+m}\sigma_1^n\sigma_2^m} e^{-\frac{1}{2\sigma_1^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right) - \frac{1}{2\sigma_2^2}\left(\sum_{j=1}^m y_j^2 - 2\mu\sum_{j=1}^m y_j + m\mu^2\right)},$$

令 $(T_1, T_2, T_3, T_4) = \left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j, \sum_{i=1}^n X_i^2, \sum_{j=1}^m Y_j^2\right)$， 有 $(t_1, t_2, t_3, t_4) = \left(\sum_{i=1}^n x_i, \sum_{j=1}^m y_j, \sum_{i=1}^n x_i^2, \sum_{j=1}^m y_j^2\right)$，

则 $p(x_1, \cdots, x_n, y_1, \cdots, y_m; \mu, \sigma_1^2, \sigma_2^2) = \frac{1}{(\sqrt{2\pi})^{n+m}\sigma_1^n\sigma_2^m} e^{-\frac{1}{2\sigma_1^2}(t_2 - 2\mu t_1 + n\mu^2) - \frac{1}{2\sigma_2^2}(t_4 - 2\mu t_3 + m\mu^2)}$，

取 $g(t_1, t_2, t_3, t_4; \mu, \sigma_1^2, \sigma_2^2) = \frac{1}{(\sqrt{2\pi})^{n+m}\sigma_1^n\sigma_2^m} e^{-\frac{1}{2\sigma_1^2}(t_2 - 2\mu t_1 + n\mu^2) - \frac{1}{2\sigma_2^2}(t_4 - 2\mu t_3 + m\mu^2)}$，

$h(x_1, \cdots, x_n, y_1, \cdots, y_m) = 1$ 与参数 $\mu, \sigma_1^2, \sigma_2^2$ 无关，

故 $(T_1, T_2, T_3, T_4) = \left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j, \sum_{i=1}^n X_i^2, \sum_{j=1}^m Y_j^2\right)$ 是参数 $(\mu, \sigma_1^2, \sigma_2^2)$ 的充分统计量.

17. 设 $\begin{pmatrix} X_i \\ Y_i \end{pmatrix}$， $i = 1, \cdots, n$ 是来自正态分布族

$$\left\{ N\left(\begin{pmatrix}\theta_1 \\ \theta_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right), \quad -\infty < \theta_1, \theta_2 < +\infty, \ \sigma_1, \sigma_2 > 0, \ |\rho| \le 1 \right\}$$

的一个二维样本，寻求$(\mu_1, \sigma_1, \mu_2, \sigma_2, \rho)$的充分统计量.

注：此题有误，应改为寻求$(\theta_1, \sigma_1, \theta_2, \sigma_2, \rho)$的充分统计量.

解：总体密度函数为

$$p(x, y; \theta_1, \sigma_1, \theta_2, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\theta_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\theta_1)(y-\theta_2)}{\sigma_1\sigma_2} + \frac{(y-\theta_2)^2}{\sigma_2^2}\right]},$$

样本联合密度函数为

$$p(x_1, y_1, \cdots, x_n, y_n; \theta_1, \sigma_1, \theta_2, \sigma_2, \rho) = \prod_{i=1}^n \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_i-\theta_1)^2}{\sigma_1^2} - 2\rho\frac{(x_i-\theta_1)(y_i-\theta_2)}{\sigma_1\sigma_2} + \frac{(y_i-\theta_2)^2}{\sigma_2^2}\right]}$$

$$= \frac{1}{(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^n} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\sigma_1^2}\sum_{i=1}^n (x_i^2 - 2\theta_1 x_i + \theta_1^2) - \frac{2\rho}{\sigma_1\sigma_2}\sum_{i=1}^n (x_i y_i - \theta_2 x_i - \theta_1 y_i + \theta_1\theta_2) + \frac{1}{\sigma_2^2}\sum_{i=1}^n (y_i^2 - 2\theta_2 y_i + \theta_2^2)\right]}$$

$$= \frac{1}{(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^n} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\sigma_1^2}\left(\sum_{i=1}^n x_i^2 - 2\theta_1\sum_{i=1}^n x_i + n\theta_1^2\right) - \frac{2\rho}{\sigma_1\sigma_2}\left(\sum_{i=1}^n x_i y_i - \theta_2\sum_{i=1}^n x_i - \theta_1\sum_{i=1}^n y_i + n\theta_1\theta_2\right) + \frac{1}{\sigma_2^2}\left(\sum_{i=1}^n y_i^2 - 2\theta_2\sum_{i=1}^n y_i + n\theta_2^2\right)\right]},$$

令 $(T_1, T_2, T_3, T_4, T_5) = \left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i, \sum_{i=1}^n X_i^2, \sum_{i=1}^n Y_i^2, \sum_{i=1}^n X_i Y_i\right)$，

有 $(t_1, t_2, t_3, t_4, t_5) = \left(\sum_{i=1}^n x_i, \sum_{i=1}^n y_i, \sum_{i=1}^n x_i^2, \sum_{i=1}^n y_i^2, \sum_{i=1}^n x_i y_i\right)$，

则 $p(x_1, y_1, \cdots, x_n, y_n; \theta_1, \sigma_1, \theta_2, \sigma_2, \rho)$

$$=\frac{1}{(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^n}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\sigma_1^2}(t_3-2\theta_1t_1+n\theta_1^2)-\frac{2\rho}{\sigma_1\sigma_2}(t_5-\theta_2t_1-\theta_1t_2+n\theta_1\theta_2)+\frac{1}{\sigma_2^2}(t_4-2\theta_2t_2+n\theta_2^2)\right]},$$

取 $g(t_1, t_2, t_3, t_4, t_5; \theta_1, \sigma_1, \theta_2, \sigma_2, \rho)$

$$=\frac{1}{(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^n}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\sigma_1^2}(t_3-2\theta_1t_1+n\theta_1^2)-\frac{2\rho}{\sigma_1\sigma_2}(t_5-\theta_2t_1-\theta_1t_2+n\theta_1\theta_2)+\frac{1}{\sigma_2^2}(t_4-2\theta_2t_2+n\theta_2^2)\right]},$$

$h(x_1, y_1, \cdots, x_n, y_n) = 1$ 与参数 $\theta_1, \sigma_1, \theta_2, \sigma_2, \rho$ 无关，

故 $(T_1,T_2,T_3,T_4,T_5)=\left(\sum_{i=1}^n X_i,\sum_{i=1}^n Y_i,\sum_{i=1}^n X_i^2,\sum_{i=1}^n Y_i^2,\sum_{i=1}^n X_iY_i\right)$ 是参数 $(\theta_1, \sigma_1, \theta_2, \sigma_2, \rho)$ 的充分统计量.

18．设二维随机变量 $X=\begin{pmatrix}X_1\\X_2\end{pmatrix}$ 服从二元正态分布，其均值向量为零向量，协方差阵为

$$\begin{pmatrix}\sigma^2+r^2 & \sigma^2-r^2\\ \sigma^2-r^2 & \sigma^2+r^2\end{pmatrix},\quad \sigma>0, r>0.$$

证明：二维统计量 $T = ((X_1 + X_2)^2, (X_1 - X_2)^2)$ 是该二元正态分布族的充分统计量.

注：此题有误，应改为 $T=\left(\sum_{i=1}^n (X_{1i}+X_{2i})^2,\sum_{i=1}^n (X_{1i}-X_{2i})^2\right)$.

证：因二元正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 的均值向量为 $\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix}$，协方差阵为 $\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$，

则 $\mu_1 = \mu_2 = 0$，$\sigma_1^2=\sigma_2^2=\sigma^2+r^2$，$\rho\sigma_1\sigma_2 = \sigma^2 - r^2$，有 $\rho=\dfrac{\sigma^2-r^2}{\sigma^2+r^2}$，$1-\rho^2=\dfrac{4\sigma^2r^2}{(\sigma^2+r^2)^2}$，

可得

$$-\frac{1}{2(1-\rho)^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]$$

$$=-\frac{1}{2}\frac{(\sigma^2+r^2)^2}{4\sigma^2r^2}\left(\frac{x_1^2}{\sigma^2+r^2}-2\frac{\sigma^2-r^2}{\sigma^2+r^2}\cdot\frac{x_1x_2}{\sigma^2+r^2}+\frac{x_2^2}{\sigma^2+r^2}\right)$$

$$=-\frac{1}{8\sigma^2r^2}[(\sigma^2+r^2)x_1^2-2(\sigma^2-r^2)x_1x_2+(\sigma^2+r^2)x_2^2]$$

$$=-\frac{1}{8\sigma^2r^2}[\sigma^2(x_1-x_2)^2+r^2(x_1+x_2)^2],$$

即总体密度函数为

$$p(x_1,x_2;\sigma,r)=\frac{1}{4\pi\sigma r}\mathrm{e}^{-\frac{1}{8\sigma^2r^2}[\sigma^2(x_1-x_2)^2+r^2(x_1+x_2)^2]},$$

样本联合密度函数为

$$p(x_{11},x_{21},\cdots,x_{1n},x_{2n};\sigma,r)=\prod_{i=1}^n\frac{1}{4\pi\sigma r}\mathrm{e}^{-\frac{1}{8\sigma^2r^2}[\sigma^2(x_{1i}-x_{2i})^2+r^2(x_{1i}+x_{2i})^2]}$$

$$=\frac{1}{(4\pi\sigma r)^n}e^{-\frac{1}{8\sigma^2r^2}\left[\sigma^2\sum_{i=1}^n(x_{1i}-x_{2i})^2+r^2\sum_{i=1}^n(x_{1i}+x_{2i})^2\right]},$$

令 $T=\left(\sum_{i=1}^n(X_{1i}+X_{2i})^2,\sum_{i=1}^n(X_{1i}-X_{2i})^2\right)$，有 $t=(t_1,t_2)=\left(\sum_{i=1}^n(x_{1i}+x_{2i})^2,\sum_{i=1}^n(x_{1i}-x_{2i})^2\right)$，

则 $p(x_{11},x_{21},\cdots,x_{1n},x_{2n};\sigma,r)=\frac{1}{(4\pi\sigma r)^n}e^{-\frac{1}{8\sigma^2r^2}(\sigma^2t_2+r^2t_1)}$，

取 $g(t_1,t_2;\sigma,r)=\frac{1}{(4\pi\sigma r)^n}e^{-\frac{1}{8\sigma^2r^2}(\sigma^2t_2+r^2t_1)}$，$h(x_{11},x_{21},\cdots,x_{1n},x_{2n})=1$ 与参数$\sigma,r$无关，

故 $T=\left(\sum_{i=1}^n(X_{1i}+X_{2i})^2,\sum_{i=1}^n(X_{1i}-X_{2i})^2\right)$是参数$(\sigma,r)$的充分统计量．

19．设 $X_1,\cdots,X_n$ 是来自两参数指数分布 $p(x;\theta,\mu)=\frac{1}{\theta}e^{-(x-\mu)/\theta}$，$x>\mu$，$\theta>0$ 的样本，证明 $(\bar{x},x_{(1)})$ 是充分统计量．

解：样本联合密度函数

$$p(x_1,x_2,\cdots,x_n;\theta,\mu)=\prod_{i=1}^n\frac{1}{\theta}e^{-\frac{x_i-\mu}{\theta}}\mathrm{I}_{x_i>\mu}=\frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^n x_i-n\mu}{\theta}}\mathrm{I}_{x_1,x_2,\cdots,x_n>\mu}=\frac{1}{\theta^n}e^{-\frac{n\bar{x}-n\mu}{\theta}}\mathrm{I}_{x_{(1)}>\mu},$$

令 $(T_1,T_2)=(\bar{X},X_{(1)})$，有 $(t_1,t_2)=(\bar{x},x_{(1)})$，即 $p(x_1,x_2,\cdots,x_n;\theta,\mu)=\frac{1}{\theta^n}e^{-\frac{nt_1-n\mu}{\theta}}\mathrm{I}_{t_2>\mu}$，

取 $g(t_1,t_2;\theta,\mu)=\frac{1}{\theta^n}e^{-\frac{nt_1-n\mu}{\theta}}I_{t_2>\mu}$，$h(x_1,x_2,\cdots,x_n)=1$ 与参数$\theta,\mu$无关，

故根据因子分解定理知 $(T_1,T_2)=(\bar{X},X_{(1)})$ 是参数$(\theta,\mu)$的充分统计量．

20．设 $Y_i\sim N(\beta_0+\beta_1x_i,\sigma^2)$，$i=1,\cdots,n$，诸 $Y_i$ 独立，$x_1,\cdots,x_n$ 是已知常数，证明 $\left(\sum_{i=1}^nY_i,\sum_{i=1}^nx_iY_i,\sum_{i=1}^nY_i^2\right)$ 是充分统计量．

解：联合密度函数

$$p(y_1,y_2,\cdots,y_n;\beta_0,\beta_1,\sigma^2)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y_i-\beta_0-\beta_1x_i)^2}{2\sigma^2}}=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n(y_i-\beta_0-\beta_1x_i)^2}$$

$$=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^ny_i^2-2\beta_0\sum_{i=1}^ny_i-2\beta_1\sum_{i=1}^nx_iy_i+\sum_{i=1}^n(\beta_0+\beta_1x_i)^2\right]},$$

令 $(T_1,T_2,T_3)=(\sum_{i=1}^nY_i,\sum_{i=1}^nx_iY_i,\sum_{i=1}^nY_i^2)$，有 $(t_1,t_2,t_3)=(\sum_{i=1}^ny_i,\sum_{i=1}^nx_iy_i,\sum_{i=1}^ny_i^2)$，

则 $p(y_1,y_2,\cdots,y_n;\beta_0,\beta_1,\sigma^2)=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\left[t_3-2\beta_0t_1-2\beta_1t_2+\sum_{i=1}^n(\beta_0+\beta_1x_i)^2\right]}$，

取 $g(T_1, T_2, T_3; \beta_0, \beta_1, \sigma^2) = \dfrac{1}{(\sqrt{2\pi}\sigma)^n} \mathrm{e}^{-\frac{1}{2\sigma^2}\left[t_3 - 2\beta_0 t_1 - 2\beta_1 t_2 + \sum\limits_{i=1}^{n}(\beta_0 + \beta_1 x_i)^2\right]}$，

$h(y_1, y_2, \cdots, y_n) = 1$ 与参数 $\beta_0, \beta_1, \sigma^2$ 无关，

故根据因子分解定理知 $(T_1, T_2, T_3) = (\sum\limits_{i=1}^{n} Y_i, \sum\limits_{i=1}^{n} x_i Y_i, \sum\limits_{i=1}^{n} Y_i^2)$ 是参数$(\beta_0, \beta_1, \sigma^2)$的充分统计量.