1、设 $p^* = \underset{p\in P_1}{\arg\min}\, D(p\|Q)$

$$r^* = \underset{p\in P_1\cap P_2}{\arg\min}\, D(p\|Q)$$

有 $p^*(x) = c_1 q(x)\exp\{\sum_{i=1}^{k}\lambda_i h_i(x)\}$

$$r^*(x) = c_2 q(x)\exp\{\sum_{i=1}^{k}\lambda_i' h_i(x) + \sum_{j=1}^{s} v_j' g_j(x)\}$$

将 p^* 投影到 $P_1\cap P_2$ 上，得到 p^{**}

$$p^{**} = \underset{\substack{p\in P_1\cap P_2\\ q\in Q}}{\arg\min}\, D(p\|q)$$

$$= c_3 p^*(x)\exp\{v_i g_i(x)\}$$

$$= c_3 c_1 q(x)\exp\{v_i g_i(x) + \lambda_i h_i(x)\}$$

$$= \tilde{c}\, q(x)\exp\{v_i g_i(x) + \lambda_i h_i(x)\}$$

则必有 $\tilde{c} = c_2$，即 p^{**} 与 r^* 分布相同、

$$\cancel{r^* = p^{**},\ \forall}$$

$$r^*(x) = p^{**}(x),\ \forall x \in \mathcal{X}$$

将Q先投影到 P_1 再到 $P_1\cap P_2$ 与直接投影到 $P_1\cap P_2$ 相同

2、(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{N}X_i$

$E\overline{X}=\frac{1}{n}E\sum_{i=1}^{N}X_i=\frac{1}{n}\sum_{i=1}^{N}EX_i=\frac{1}{n}\times n\mu=\mu$

即 $\overline{X}$ ~~无偏~~ 是 μ 的无偏估计

(2) $\cancel{E\left[S_{n-1}^2\times(n-1)\right]}$

$=\cancel{n-1}\ \cancel{E\left(\sum_{i=1}^{n}(X_i-\overline{X})^2\right)}$

$=E\left\{\sum_{i=1}^{n}X_i^2+n\overline{X}^2-2\sum_{i=1}^{n}\overline{X}X_i\right\}$

$(n-1)ES_{n-1}^2$

$=\cancel{(n}\ E\left\{\sum_{i=1}^{n}(X_i-\mu+\mu-\overline{X})^2\right\}$

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

$=\sum_{i=1}^{n}E(X_i-\mu)^2+nE(\overline{X}-\mu)^2\cancel{+2(\mu-\overline{X})}-2nE(\overline{X}-\mu)^2$

$=\cancel{n\,var}\sum_{i=1}^{n}varX_i\ \cancel{+n\,var\ (n-2)}-n\,var\overline{X}$

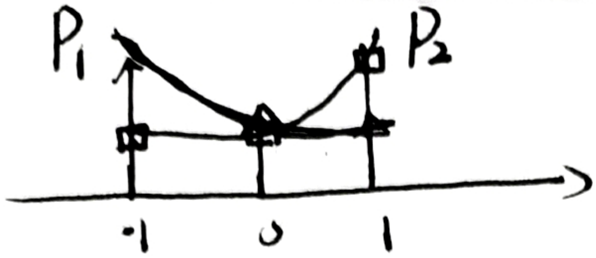
$=n\sigma^2-n\cdot\frac{1}{n^2}\sum_{i=1}^{n}varX_i$

$=(n-1)\sigma^2$

则 $ES_{n-1}^2=\sigma^2$ 无偏. $ES_n^2=E\frac{n-1}{n}S_{n-1}^2=\frac{n-1}{n}\sigma^2$

有偏

3、



$$D(P_2\|P_1)=\frac{1}{4}\log\frac{\frac{1}{4}}{\frac{1}{2}}+\frac{1}{4}\log\frac{\frac{1}{4}}{\frac{1}{4}}+\frac{1}{2}\log\frac{\frac{1}{2}}{\frac{1}{4}}$$

$$=0.25$$

最佳误差指数为 0.25

4、~~设 Pr{Y=k}=~~

设 $Pr\{Y=k\}=r_k$，则该分布与题中分布相对熵为 $D(r\|p)=\sum r_i\log\frac{r_i}{p^{i-1}(1-p)}$

问题转化为 $\min D(r\|p)$

$$\text{s.t.}\begin{cases}\sum r_i=1\\ \sum i r_i=\alpha\end{cases}$$

设 $J(r)=\sum r_i\log\frac{r_i}{p^{i-1}(1-p)}+\lambda_1\sum r_i+\lambda_2\sum i r_i$

$$\frac{\partial J}{\partial r_i}=\log r_i-\log(p^{i-1}(1-p))+\lambda_1+\lambda_2 i=0.$$

得 $r_i=p^{i-1}(1-p)c_1c_2^{i}$

~~如下~~ 显然，r_i 也~~最大~~服从几何分布，且 $p=1-\frac{1}{\alpha}$ 即

$$r_i=\left(1-\frac{1}{\alpha}\right)^{i-1}\frac{1}{\alpha}$$

则与 $\{x_i\}$ 相对熵 $D(r\|p)=\log\frac{-p}{(1-p)(1-\alpha)}+\alpha\log\frac{\alpha-1}{\alpha p}$

(1) ~~$\Pr\{\frac{1}{n}\sum x_i\geq\alpha\}=$~~

$$-\frac{1}{n}\log\Pr\left\{\frac{1}{n}\sum x_i\geq\alpha\right\}\cancel{(n\to\infty)}=D(r\|p).$$

则 $\Pr\left\{\frac{1}{n}\sum x_i\geq\alpha\right\}=\cancel{\exp}\,2^{-n\log\frac{-p}{(1-p)(1-\alpha)}-n\alpha\log\frac{\alpha-1}{\alpha p}}$

(2) $\Pr\left\{X_1=k\,\middle|\,\frac{1}{n}\sum x_i\geq\alpha\right\}=r_k=\left(1-\frac{1}{\alpha}\right)^{i-1}\frac{1}{\alpha}$

(3) $p=\frac{1}{2}$, $\alpha=4$ 时

$$-\frac{1}{n}\log\Pr\left\{\frac{1}{n}\sum x_i\geq\alpha\right\}=D(r\|p)=\log 27-\log 16=0.755$$

$$\Pr\left\{X_1=k\,\middle|\,\frac{1}{n}\sum x_i\geq\alpha\right\}=r_k=\left(\frac{3}{4}\right)^{k-1}\frac{1}{4}$$