概率论与数理统计 第十五次作业参考题解

2023.12.26

Q1

(1) 最小二乘法的结果为 $\hat{\beta}_1=\frac{\sum_{i=1}^n(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^n(x_i-\bar{x})^2},\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}$，故 $\hat{\beta}_0+\hat{\beta}_1\bar{x}=\bar{y}$，即最小二乘法拟合的直线经过点 $(\bar{x},\bar{y})$。

(2) $\mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1)=\mathrm{Cov}(\bar{y}-\hat{\beta}_1\bar{x},\hat{\beta}_1)=\mathrm{Cov}(\bar{y},\hat{\beta}_1)-\bar{x}\mathrm{Var}(\hat{\beta}_1)$，其中 $\mathrm{Cov}(\bar{y},\hat{\beta}_1)=\mathrm{E}(\bar{y}\hat{\beta}_1)-\mathrm{E}(\bar{y})\mathrm{E}(\hat{\beta}_1)=\mathrm{E}(\frac{\bar{y}\sum_{i=1}^n(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^n(x_i-\bar{x})^2})-\beta_1(\beta_0+\beta_1\bar{x})=\frac{\sum_{i=1}^n(x_i-\bar{x})\mathrm{E}((y_i-\bar{y})\bar{y})}{\sum_{i=1}^n(x_i-\bar{x})^2}-\beta_1(\beta_0+\beta_1\bar{x})$，利用 $\mathrm{E}(y_iy_j)=\mathrm{E}(y_i)\mathrm{E}(y_j)=(\beta_0+\beta_1x_i)(\beta_0+\beta_1x_j)(i\neq j),\mathrm{E}(y_i^2)=\sigma^2+(\beta_0+\beta_1x_i)^2$，得 $\mathrm{E}((y_i-\bar{y})\bar{y})=\frac{\sum_{j=1}^n\mathrm{E}(y_iy_j)}{n}-(\frac{\sigma^2}{n}+(\beta_0+\beta_1\bar{x})^2)=\frac{\sigma^2+\sum_{j=1}^n(\beta_0+\beta_1x_i)(\beta_0+\beta_1x_j)}{n}-(\frac{\sigma^2}{n}+(\beta_0+\beta_1\bar{x})^2)=\frac{\sum_{j=1}^n(\beta_0+\beta_1x_i)(\beta_0+\beta_1x_j)}{n}-(\beta_0+\beta_1\bar{x})^2=\beta_1(x_i-\bar{x})(\beta_0+\beta_1\bar{x})$，代入得 $\mathrm{Cov}(\bar{y},\hat{\beta}_1)=0$，故 $\mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1)=-\bar{x}\mathrm{Var}(\hat{\beta}_1)=-\frac{\bar{x}\sigma^2}{S_{xx}}$。当 $\bar{x}=0$ 时，$\mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1)=0$，此时 $\hat{\beta}_0,\hat{\beta}_1$ 不相关。

(3) 即最大化 S_{xx}。若 n 为偶数，则一半 x_i 取 1，另一半 x_i 取 -1；若 n 为奇数，则 $\frac{n-1}{2}$ 个 x_i 取 1，其他 x_i 取 -1，或反之。具体证明过程参见 https://cloud.tsinghua.edu.cn/f/94fc5b491af0419b8085/，感谢计 23 班秦晨阳同学提供解答。也可采用多元微积分的方法，感兴趣的同学可自行尝试。

(4) 相应的最小二乘估计为 $\hat{\beta}_1=\frac{\sum_{i=1}^n x_iy_i}{\sum_{i=1}^n x_i^2}$，此时相应的 $\hat{\sigma}^2=\frac{\sum_{i=1}^n(y_i-\hat{\beta}_1x_i)^2}{n-1}$。记 $\hat{y}_0=\hat{\beta}_1x_0$，则可以证明 $\frac{y_0-\hat{y}_0}{\sigma\sqrt{1+\frac{x_0^2}{\sum_{i=1}^n x_i^2}}}\sim N(0,1),\frac{(n-1)\hat{\sigma}^2}{\sigma^2}\sim\chi^2(n-1)$ 且二者独立，故 $\frac{y_0-\hat{y}_0}{\hat{\sigma}\sqrt{1+\frac{x_0^2}{\sum_{i=1}^n x_i^2}}}\sim t(n-1)$，据此给出 y_0 的 $(1-\alpha)$-置信的区间估计为 $[\hat{y}_0-t_{\frac{\alpha}{2}}(n-1)\hat{\sigma}\sqrt{1+\frac{x_0^2}{\sum_{i=1}^n x_i^2}},\hat{y}_0+t_{\frac{\alpha}{2}}(n-1)\hat{\sigma}\sqrt{1+\frac{x_0^2}{\sum_{i=1}^n x_i^2}}]$。

Q2

(1) 直接优化对数似然函数即得 $(\sigma^2)^*=SSE/n$。

(2) 参考 https://math.stackexchange.com/questions/3319241/proof-that-ess-e-n-2-sigma2。