- $y = f(x) + e$, $f(x) = a + bx$.
- 最小二乘拟合系数、最大似然估计、最优线性无偏估计:
  $\hat{a} = \bar{y} - \hat{b}\bar{x}$, $\quad \hat{b} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\ell_{xy}}{\ell_{xx}}$.
- 残差平方和: $Q = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=3}^{n} Z_i^2$,
  回归平方和: $U = \sum_i(\hat{y}_i - \bar{y})^2 = (\sqrt{\ell_{xx}}b + Z_2)^2$,
- $\frac{1}{\sigma^2}Q \sim \chi^2(n-2)$, $\quad \frac{1}{\sigma^2}U_0 \sim \chi^2(1)$, 在某种意义下, $U_0 \leqslant U_b$.
- Q 与U 相互独立.
- 直观: $Q, U, \ell_{yy} = \sum_i(y_i - \bar{y})^2$ 都是$N(0,*)$ 的平方和.
  ℓ_{yy}: 自由度为$n-1$ (1个约束条件$\sum_i(y_i - \bar{y}) = 0$).
  U: 自由度(维数)为1, $\hat{y}_i - \bar{y} = \hat{b}(x_i - \bar{x})$.
  $Q = \ell_{yy} - U$: 自由度为$(n-1) - 1 = n-2$.

假设检验问题$H_0: b=0 \leftrightarrow H_1: b\neq 0$.

- 否定H_0, 则表明y 与x 之间有线性依赖关系.
- 若H_0 成立, 则$\frac{1}{\sigma^2}Q \sim \chi^2(n-2)$, $\frac{1}{\sigma^2}U \sim \chi^2(1)$, $\frac{U}{Q/(n-2)} \sim F(1, n-2)$.
- 否定域: $W=\{(\vec{x},\vec{y}): \frac{U}{Q/(n-2)} > \lambda\}$.
  在某种意义下, $U_0 \leqslant U_b$.
- 根据水平α 选择λ. $P(F_{1,n-2}>\lambda)=\alpha$.

例2.1. 根据散点图建立函数关系$y = e^a e^{bx}$,

即回归关系式$z = \log y = a + bx + e$.

$x =$ 注射后天数, $y =$ 金残留量.

- $n = 7$, 数据$(x_i, \quad y_i \to z_i = \log y_i)$, $i = 1, \cdots, 7$.
- 求$\bar{x}$, $\bar{z}$; $\hat{b} = \frac{\ell_{xz}}{\ell_{xx}}$, $\hat{a} = \bar{z} - \hat{b}\bar{x}$.
  以及$\hat{z}_i = \hat{a} + \hat{b}x_i$.
- 求残差平方和: $Q = \sum_i (z_i - \hat{z}_i)^2$
  与回归平方和: $U = \sum_i (\hat{z}_i - \bar{z})^2$.
- 根据$P(F_{1,n-2} = F_{1,5} > \lambda) = \alpha = 0.05$, 查表得$\lambda = 6.61$.
  $\frac{U}{Q/(n-2)} = 344.82 > \lambda$ 否定H_0, 强烈认可z 线性依赖于x.
- 自习图9.9.2 $\sim$ 9.2.5.

§9.3 多元线性回归

- 函数关系: $y = a + b_1x_1 + \cdots + b_px_p + e = a + \mathbf{x}\mathbf{b} + e$.
  $\mathbf{x} = (x_1, \cdots, x_p)$ 行向量, $\mathbf{b} = (b_1, \cdots, b_p)^T$ 列向量.
- 数据: y_i; $\mathbf{x}_i = (x_{i1}, \cdots, x_{ip})$, $i = 1, \cdots, n$.
  均值: $\bar{y}$; $\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i$.
- 回归模型: $y_i = a + \mathbf{x}_i\mathbf{b} + e_i$, $e_1, \cdots, e_n \sim$ i.i.d. $N(0, \sigma^2)$.
  $\mathbf{y} = a\vec{1} + \mathbf{X}\mathbf{b} + \mathbf{e}$, 其中$\mathbf{y} = (y_i)_{n\times 1}$, $\mathbf{X} = (x_{ij})_{n\times p}$.
- 最小二乘、最优(定理3.1,3.3):
  $\hat{a} = \bar{y} - \hat{b}\bar{x} \longrightarrow \bar{y} - \mathbf{X}\hat{\mathbf{b}}$,
  $\hat{\mathbf{b}} = \frac{1}{\ell_{xx}}\ell_{xy} \longrightarrow (\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T\tilde{\mathbf{y}} = (\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T\mathbf{y}$,
  $\tilde{\mathbf{x}}_i = \mathbf{x}_i - \bar{\mathbf{x}}$, $\tilde{\mathbf{y}} = \mathbf{y} - \bar{y}\mathbf{1}$. $\underline{\tilde{\mathbf{X}}^T\mathbf{1} = 0}$.
- $\hat{\mathbf{b}} = (\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T\mathbf{X}\mathbf{b} + (\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T\mathbf{e}$

假设检验$H_0: \mathbf{b}=0 \leftrightarrow H_1: \mathbf{b}\neq 0.$

- $\hat{y}_i - \bar{y} = \hat{b}(x_i - \bar{x}) \longrightarrow (\mathbf{x}_i - \bar{\mathbf{x}})\hat{\mathbf{b}}$
  $U = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n ((\mathbf{x}_i - \bar{\mathbf{x}})\hat{\mathbf{b}})^2 = \|\tilde{\mathbf{X}}\hat{\mathbf{b}}\|^2.$
  自由度(维数)为p: $\hat{y}_i - \bar{y} = \sum_{j=1}^p \hat{b}_j (x_{ij} - (\bar{\mathbf{x}})_j).$
- $\frac{1}{\sigma^2} U \overset{H_0}{\sim} \chi^2(p).$
  $\tilde{\mathbf{X}}\hat{\mathbf{b}} \overset{H_0}{=} \underline{\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T}\mathbf{e}.$
  $\frac{1}{\sigma^2} EU = \mathrm{Tr}(\underline{\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T}\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T) =$
  $\mathrm{Tr}(\underline{(\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^T}\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^T\tilde{\mathbf{X}})^{-1}_{p\times p}(\tilde{\mathbf{X}}^T\underline{\tilde{\mathbf{X}}})_{p\times p}) = \mathrm{Tr}(\mathbf{I})_{\mathbf{p}\times\mathbf{p}} = \mathbf{p}.$
- $\frac{1}{\sigma^2} Q \sim \chi^2(n-1-p).$
  $y_i - \hat{y}_i = (y_i - \bar{y}) - (\hat{y}_i - \bar{y}) = \tilde{y}_i - (\mathbf{x}_i - \bar{\mathbf{x}})\hat{\mathbf{b}}.$
  $Q = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \|\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\hat{\mathbf{b}}\|^2.$
- $\frac{U/p}{Q/(n-p-1)} \overset{H_0}{\sim} F(p, n-p-1).$
  否定域: $W = \{\frac{U/p}{Q/(n-p-1)} > \lambda\}$, 其中$P(F_{p,n-p-1} > \lambda) = \alpha.$

§10.1 统计决策问题概述

例1.1. θ_1好, θ_2坏; a_1保留, a_2更换, a_3修理.

表 10.1.1　损失函数 $L(\theta,a)$的值

$L(\theta,a)$ \ a ; θ	a_1	a_2	a_3
θ_1	0	10	5
θ_2	12	1	6

- 状态: $\theta = \theta_1$ 或θ_2.
- 行动: a_1, a_2, a_3. 行动空间$A = \{a_1, a_2, a_3\}$.
- 损失: 比如$\underline{L(\theta_1, a_2) = 10}$: 如果是好零件, 更换它, 损失为10.
- 本应根据θ 采取行动使得L 最小, 但是θ 不可观测!
- 样本: $X = 1$(正常); $X = 0$(发烫). 可观测. 根据x 采取行动.
- X 与θ 有关: $X \sim P_\theta$.

决策与风险.

- 决策函数: 根据观测值采取行动.

  如: δ_7 = 若发烫则a_3 = 修理, 若正常则a_1 = 保留.

  表 10.1.3 9个决策函数的列表

δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
$\delta(0)$	a_1	a_1	a_1	a_2	a_2	a_2	a_3	a_3	a_3
$\delta(1)$	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3

  每个X 的观测值x, 有3种可选行动. 因此, 共3×3 个决策.

- 风险函数: 决策带来的平均损失.

  P_θ: $P_{\theta_1=\text{好}}(X=0=\text{发烫})=0.3$, $P_{\theta_2}(X=0)=0.6$.

  如: $R(\theta_1,\delta_2)=P_{\theta_1}(X=0)L(\theta_1,a_1)+P_{\theta_1}(X=1)L(\theta_1,a_2)$

  $=0.3\times 0+0.7\times 10=7.$

决策的优劣:

表 10.1.4 各决策函数的风险函数值

δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
$R(\theta_1,\delta)$	0	7	3.5	3	10	6.5	1.5	8.5	5
$R(\theta_2,\delta)$	12	7.6	9.6	5.4	1	3	8.4	4	6

- 不可容许的: 如δ_2.

  $R(\theta,\delta_2) \geqslant R(\theta,\delta_4)$, $\forall\theta$, 且存在严格不等号.

- 该例题不存在一致最优的δ: 对任意$\tilde{\delta}$ 都有

  $R(\theta,\tilde{\delta}) \geqslant R(\theta,\delta)$, $\forall\theta$.

- 极小极大准则: 选δ_4 使得其最大风险5.4 是最小的.

贝叶斯决策.

- 将θ 视为随机的: 如, 假设某零件“好”的概率为0.7, 先验分布: $\pi(\theta_1) = 0.7$, $\pi(\theta_2) = 0.3$.
- 平均风险: $\pi(\theta_1)R(\theta_1, \delta) + \pi(\theta_2)R(\theta_2, \delta)$.

表 10.1.5　各决策函数的平均风险值

δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
平均风险	3.6	5.13	12.48	3.72	7.20	5.45	3.57	7.15	5.3

- 选择δ_7, 使得平均风险达到最小.

一般情形.

- 状态(空间): $\theta \in \Theta$.
  行动(空间): $a \in A$.
  损失函数: $L(\theta, a)$.
- 样本(空间): $X \sim P_\theta$, 取值$x \in \mathcal{X}$.
  策略函数: $\delta : \mathcal{X} \to A$, $\delta(x) = a$.
  风险函数: $R(\theta, \delta) = E_\theta L(\theta, \delta(X))$, $a = \delta(X)$.
- 极小极大准则:
  选择δ^* 使得$\max_{\theta\in\Theta} R(\theta, \delta)$ 在$\delta = \delta^*$ 达到最小.
- 贝叶斯决策:
  将θ 视为取值于Θ 的随机变量, 分布为π.
  选择δ^* 使得$E_\pi R(\theta, \delta)$ 在$\delta = \delta^*$ 达到最小.