矩估计

- 均值: $E(X) = \bar{X}$
- 方差: $\frac{1}{n}\sum X_i^2 = DX + (EX)^2$

极大似然估计

- 似然函数: $L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n f(x_i, \theta) \Rightarrow$ 最大
- 无偏性: $E[\hat{\theta}(x_1, \ldots, x_n)] = \theta$
- 相合性:
  - **概率收敛**: $\lim_{n\to\infty} P[|\hat{\theta}_n - \theta| \le \varepsilon] = 1$，或
  - **均方收敛**: $\lim_{n\to\infty} E(\hat{\theta}_n) = E(\theta_n)$ & $\lim_{n\to\infty} D\hat{\theta}_n = 0$

C-R下界

- **信息量**: $I(\theta) = E[(\frac{\partial \ln f(x,\theta)}{\partial\theta})^2] = -E[\frac{\partial^2 \ln f(x,\theta)}{\partial\theta^2}]$ （第二个仅正则条件满足可用，均匀分布不可）
- **C-R下界**:
  - 一般形式: $D(T) \ge \frac{[g'(\theta)]^2}{nI(\theta)}$ (其中 $T = g(\hat{\theta})$ 是 $\tau = g(\theta)$ 的无偏估计)
  - 特殊形式: $D(\hat{\theta}) \ge \frac{1}{nI(\theta)}$ (对 θ 的无偏估计)

渐近分布与假设检验

分布	统计量	假设条件	拒绝域
正态	$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$	$H_0: \mu = \mu_0, \sigma^2$已知	$\|Z\| > z_{1-\alpha/2}$
t	$t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \sim t(n-1)$	$H_0: \mu = \mu_0, \sigma^2$未知	$\|t\| > t_{1-\alpha/2}(n-1)$
卡方	$\chi^2 = \frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma_0^2} \sim \chi^2(n)$	$H_0: \sigma^2 = \sigma_0^2$, μ已知	$\chi^2 < \chi^2_{\alpha/2}(n)$ 或 $\chi^2 > \chi^2_{1-\alpha/2}(n)$
卡方	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(n-1)$	$H_0: \sigma^2 = \sigma_0^2$, μ未知	$\chi^2 < \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 > \chi^2_{1-\alpha/2}(n-1)$

期望置信区间

- 方差已知: $\frac{(\bar{X}-\bar{Y})-(\mu_x-\mu_y)}{\sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}}} \sim N(0,1)$
- 方差未知且相等: $\frac{(\bar{X}-\bar{Y})-(\mu_x-\mu_y)}{S_w\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}} \sim t(n_x+n_y-2)$

  其中 $S_w^2 = \frac{(n_x-1)S_x^2+(n_y-1)S_y^2}{n_x+n_y-2}$

- 方差未知且不相等但样本大小相等 (大小不等则n为较小值)：

  定义 $Z_i = X_i - Y_i$，其中：

  $\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2$

  统计量：$T = \frac{\bar{Z}}{S/\sqrt{n}} \sim t(n-1)$

方差比检验

$H_0 : \sigma_x^2 = \sigma_y^2$

$\frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} \sim F(n_x - 1, n_y - 1)$

拒绝域: $F > F_{1-\alpha/2}(n_x - 1, n_y - 1)$ 或 $F < F_{\alpha/2}(n_x - 1, n_y - 1)$

回归分析

- 最小二乘回归:
  $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$

  $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$
- F检验:
  $H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0$
  $F = \frac{U/k}{Q_e/(n-k-1)} = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k}$
  拒绝域: $F > F_{1-\alpha}(k, n-k-1)$

  其中:
  $R^2 = \frac{U}{S_{yy}}$

  $U = \hat{\beta}_1 S_{xy}$

  $Q_e = S_{yy} - U$
- $\hat{y}$ 的预测区间估计:
  $\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} + t_{1-\alpha/2}(n-2)$
  $= \sqrt{\frac{Q_e}{n-2}} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} \cdot t_{1-\alpha/2}(n-2)$

方差分析表

变异来源	自由度	离差平方和	F
因素间	$a-1$	$S_A = \sum_{i=1}^{a} n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$	$F = \frac{S_A/(a-1)}{S_E/(n-a)} \sim F(a-1, n-a)$
误差	$n-a$	$S_E = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i\cdot})^2$	
总和	$n-1$	$S_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{\cdot\cdot})^2$	

方差分析的假设检验:

- H_0: 各组均值相等
- 拒绝域: $F > F_{1-\alpha}(a-1, n-a)$

拟合优度检验

$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \sim \chi^2(n-s-1)$

- H_0: 样本来自指定的分布
- 拒绝域: $\chi^2 > \chi^2_{1-\alpha}(n-s-1)$

注: 当使用极大似然法估计的参数个数为s时（正态分布为2）

列联表检验

	第一类	...	第c类	行和
第一组	n_{11}	...	n_{1c}	$n_{1\cdot}$
...	...	...	...	...
第r组	n_{r1}	...	n_{rc}	$n_{r\cdot}$
列和	$n_{\cdot 1}$	...	$n_{\cdot c}$	n

- 二维列联表:
  - H_0: 行变量与列变量相互独立
  - $\chi^2 = n \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(n_{ij} - \hat{n}_{ij})^2}{n_{i\cdot} n_{\cdot j}} \sim \chi^2((r-1)(c-1))$
  - 拒绝域: $\chi^2 > \chi^2_{1-\alpha}((r-1)(c-1))$

  其中 $\hat{n}_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n}$
- 四格表: $\chi^2 = \frac{n(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)} \sim \chi^2(1)$