

Chapter 12 Simple Regression

- ① 利用方程求值
- ② Test for slope: $H_0: \beta_1 = 0$, $H_1: \beta_1 \neq 0$, p-value
- ③ 注意: $df_1 = 1$, $df_2 = n - 2$, total $df = n - 1$
- ④ confidence interval for mean of Y
prediction interval for individual Y (用 sheet)
- ⑤ confidence interval for slope (直接用 excel 结果)
- ⑥ 利用 ANOVA 表回答问题, 计算 ANOVA 表中的值. (MSR, MSE, Fcrit, p-value)
- ⑦ $R^2 = \% \text{ of variation the model explains}$
- ⑧ r (correlation coefficient) 负相关 -1 — 0 — $+1$ 正相关
- ⑨ "The standard error is too high for this model to be of any predictive use."
- ⑩ $t_{calc} = \frac{\beta_1}{S_{b_1}}$, $t_{calc} = r \sqrt{\frac{(n-2)}{1-r^2}}$

Chapter 13 Multiple Regression

- ★ ① 判断 intercept 是否有意义 (注意单位实际意义!)
- ② 注意: $df_1 = c$, $df_2 = n - c - 1$, $df = n - 1$
 - ③ 计算 ANOVA 值 (R^2 , R^2_{adj}) $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$
 - ④ 解释方程的实际意义.
 - ⑤ 区别 quantitative / categorical variable, 对于类别变量, 变量数应为 $n - 1$ (dummy)
 - ⑥ 检验: $H_0: \beta_0 / \beta_1 = 0$ (significant predictor)
 - ⑦ predictor 数量:
 - 1) Evans' Rule: 10 observations / predictor
 - 2) Doane's Rule: 5 observations / predictor.
- 太大 R^2 无法提供准确性

Chapter 14 Time-Series Analysis

① Trailing Moving Average (TMA) $\hat{y}_t = \frac{y_t + y_{t-1} + \dots + y_{t-m+1}}{m}$

• smoothes the past fluctuations

• 可用于预测 y_{t+1} , 但最好只作为 one-period-ahead ~~forecast~~ forecast

• 用 "Moving Average" 计算

② Centered Moving Average (CMA) $\hat{y}_t = \frac{y_{t-1} + y_t + y_{t+1} + \dots}{2}$

• 手动计算. (例, 对于季度: $CMA_3 = \frac{AVG(1 \sim 4) + AVG(2 \sim 5)}{2}$)

③ Seasonal Index, deseasonalization, adjusted seasonal index.

④ Forecasting is not a major factor affect capital investments.

⑤ 判断 Moving Average 是不是好方法 / useful: 看和原图的符合程度.

⑥ 区分 linear, cyclical, irregular 趋势, 并判断是否 consistent.

⑦ Exponential Smoothing: $F_{t+1} = \alpha y_t + (1-\alpha) F_t$

α : smoothing constant

$1-\alpha$: damping factor

α 越大越平滑.

default: $\alpha = 0.2$

MSE: $\sum (y_i - e_i)^2$

MAPE: $(\sum \frac{|y_i - e_i|}{y_i}) \cdot \frac{100}{n}$ (该值为百分数)

F_1 计算方法: 1) Method A: $F_1 = y_1$

2) Method B: $F_1 = \frac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}{6}$ (default)

• As $\alpha \uparrow$, the value of forecast \downarrow

⑧ 利用数据做图, 分别用 linear, exponential, quadratic

⑨ Seasonal Data regression 注意用的是原数据还是 deseasonalized data.

Chapter 15 Chi-Square Tests

① Contingency table: for a $r \times c$ table: d.f. = $(r-1)(c-1)$

$$\chi^2_{\text{calc}} = \sum \frac{(y_{ijk} - e'_{ijk})^2}{e'_{ijk}} \quad , \quad e'_{ijk} = \frac{R_j C_k}{n}$$

H_0 : Variable A is independent of B

⑤ 计算 t -stat, d.f. p -value

② GOF test 分辨 H_0, H_1 区别不同分布的 d.f.

Multinomial GOF: $H_0: \pi_1 = 0.1, \pi_2 = 0.2, \dots$ $df = c - 1$

Uniform
Poisson
Normal

H_0 : The population follows a _____ distribution

$df = c - 0 - 1$
 $df = c - 2$ (λ)
 $df = c - 3$ (σ, μ)

③ 对于 poisson, 需要将 data 放入框中.

对于 "X or more" 需满足 Cochran's Rule: $ej \geq 5$ (上下限都需满足).

• 且该处 $P(X=x) = 1 - P(X=x-1)$

- $ef = \text{POISSON.DIST}(x, \lambda, a) \times 100$

注意看题是几个框，注意看清题目选项

- 可用 = CHISQ.TEST() 验证, $w/h = 1 - \text{CHISQ.DIST}()$ 为全.