1 公式

$\mathrm{Var}(X)=\mathbb{E}(X^2)-(\mathbb{E}(X))^2$
$\Sigma_x=\mathbb{E}\left[(x-\mathbb{E}[x])(x-\mathbb{E}[x])^\top\right]$
$y=Ax\Rightarrow\Sigma_y=A\Sigma_xA^\top$
$A\perp\!\!\!\perp B\Rightarrow\sigma^2_{A\pm B}=\sigma^2_A+\sigma^2_B$
$X=\begin{bmatrix}1&1&\cdots&1\\x_1&x_2&\cdots&x_n\end{bmatrix}^\top, X^\top X=\begin{bmatrix}1&\sum x_i\\\sum x_i&\sum x_i^2\end{bmatrix}$

2 Simple Linear Regression

2.1 Model Fitting

Residual: $e_i=y_i-\hat{y}_i$
$y_i=\beta_0+\beta_1x_i+\epsilon_i,\mathbb{E}[\epsilon_i]=0,\mathrm{Var}(\epsilon_i)=\sigma^2$
$\begin{cases}S_{xx}=\sum_{i=1}^n\left(x_i-\overline{x}\right)^2\\S_{xy}=\sum_{i=1}^n\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)\\S_{yy}=\sum_{i=1}^n\left(y_i-\overline{y}\right)^2\end{cases}$
$\hat\beta_0=\overline{y}-\hat\beta_1\overline{x},\hat\beta_1=\frac{S_{xy}}{S_{xx}}$
Gauss-Markov Theorem (quality of LSE): Among all estimates that are linear combination of $y_1,\cdots,y_n$ and unbiased, the LSE has the smallest variance.
$\sum_{i=1}^n e_i=0,\sum_{i=1}^n x_ie_i=0,\sum_{i=1}^n y_i=\sum_{i=1}^n\hat{y}_i$

2.2 Statistic Inference & Model Test

Error sum of squares: $\mathrm{SSE}=\sum_{i=1}^n\left(y_i-\hat{y}_i\right)^2$
$s^2=\mathrm{MSE}=\frac{\mathrm{SSE}}{n-2}$
Regression sum of squares: $\mathrm{SSR}=\sum_{i=1}^n\left(\hat{y}_i-\overline{y}\right)^2$
$S_{yy}=\mathrm{SSR}+\mathrm{SSE}$
$\hat\beta_0=\boldsymbol{l}^\top\boldsymbol{y},\hat\beta_1=\boldsymbol{k}^\top\boldsymbol{y},k_i=\frac{x_i-\overline{x}}{S_{xx}},l_i=\frac{1}{n}-k_i\overline{x}$
$\sum_{i=1}^n l_i=0,\sum_{i=1}^n l_ix_i=0$
$\frac{\hat\beta_1-\beta_1}{\sqrt{\sigma^2/S_{xx}}}\sim\mathcal{N}(0,1),\frac{\hat\beta_0-\beta_0}{\sqrt{\left(\sigma^2\sum_{i=1}^n x_i^2\right)/(nS_{xx})}}\sim\mathcal{N}(0,1)$
$\frac{\hat\beta_1-\beta_1}{\sqrt{s^2/S_{xx}}}\sim t_{n-2},\frac{\hat\beta_0-\beta_0}{\sqrt{\left(s^2\sum_{i=1}^n x_i^2\right)/(nS_{xx})}}\sim t_{n-2}$

$\frac{\hat{y}_0-\mathbb{E}[y_0]}{s\sqrt{\frac{1}{n}+\frac{(x_0-\overline{x})^2}{S_{xx}}}}\sim t_{n-2}$
$\frac{y_{\text{new}}-\hat{y}_{\text{new}}}{s\sqrt{1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{S_{xx}}}}\sim t_{n-2}$
$\mathcal{R}^2=\frac{\mathrm{SSR}}{S_{yy}}=\frac{S_{xy}^2}{S_{xx}S_{yy}}=r_{xy}^2$
ANOVA for SLR:

Source of variation	df	Sum of squares (SS)	MS	F	p-value
Regression	1	$SSR=\sum(\hat{y}_i-\bar{y})^2$	$\mathrm{MS}_{\mathrm{Reg}}$	$\mathrm{MS}_{\mathrm{Reg}}/s^2$	
Residual	$n-2$	$SSE=\sum(y_i-\hat{y}_i)^2$	s^2		
Total	$n-1$	$S_{yy}=\sum(y_i-\bar{y})^2$			

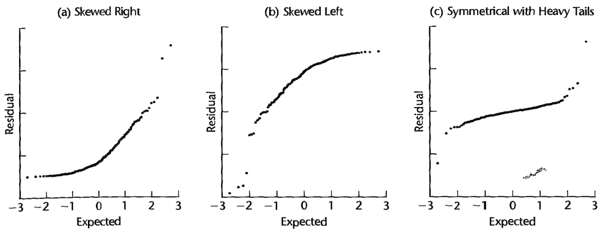
$\mathrm{MS}=\frac{\mathrm{SS}}{\mathrm{df}}$
$\mathrm{SSR}\sim\sigma^2\chi_1^2,\mathrm{SSE}\sim\sigma^2\chi_{n-2}^2$
$F=\frac{\mathrm{MS_{reg}}}{s^2}\sim F(1,n-2)$
$\mathbb{E}[\mathrm{MS_{reg}}]=\sigma^2+\beta_1^2S_{xx}$
F 越大说明 $\mathrm{MS_{reg}}$ 越大，也就是 SSR 的贡献更大，更能说明 $\beta_1\neq0$, X effective in explaining variation in Y

2.3 Model Diagnostics

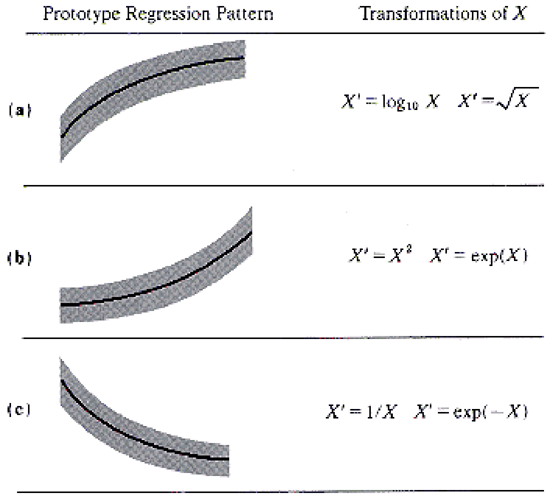
Standardized (semistudentized) residual: $e_i^*=\frac{e_i}{\sqrt{\mathrm{MSE}}}$
Rule of Thumb: $|e_i^*|>3\Leftrightarrow$ outliers
QQ-Plot: k-th smallest, $\sqrt{\mathrm{MSE}}\left[z\left(\frac{k-0.375}{0.25}\right)\right]$

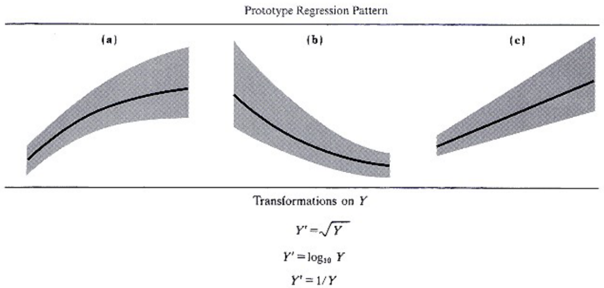
residual	expected residual
$e_{\min}$	$\sqrt{\mathrm{MSE}}\left[z\left(\frac{1-0.375}{n+0.25}\right)\right]$
$e_{\text{2nd smallest}}$	$\sqrt{\mathrm{MSE}}\left[z\left(\frac{2-0.375}{n+0.25}\right)\right]$
$\vdots$	$\vdots$
$e_{\max}$	$\sqrt{\mathrm{MSE}}\left[z\left(\frac{n-0.375}{n+0.25}\right)\right]$

FIGURE 3.9 Normal Probability Plots when Error Term Distribution Is Not Normal.



左边：往下尾巴大；右边：往上尾巴大





3 Calculator

$\sigma^2\mathtt{x}=\frac{1}{n}\sum_{i=1}^n(x_i-\overline{x})^2$
$\mathtt{s}^2\mathtt{x}=\sqrt{\frac{1}{n-1}\sum_{i=1}^n(x_i-\overline{x})^2}$
$S_{xx}=n\cdot\sigma^2\mathtt{x},S_{yy}=n\cdot\sigma^2\mathtt{y}$
$S_{xy}=\Sigma\mathtt{xy}-n\cdot\overline{\mathtt{x}}\cdot\overline{\mathtt{y}}$