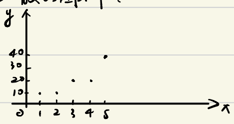
2.14

解：(1) 散点图如下：



(2) 由图像：

大致有线性关系

(3) 由题可知：$L_{xx}=\sum(x_i-\bar{x})^2=10$，$L_{yy}=\sum(y_i-\bar{y})^2=600$

$$L_{xy}=\sum(x_i-\bar{x})(y_i-\bar{y})=70$$

令回归方程为 $\hat{y}_i=\hat{\beta}_0+\hat{\beta}_1x_i$，其中，

$$\hat{\beta}_1=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}=7$$

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=-1$$

$$\Rightarrow \hat{y}_i=-1+7x_i$$

(4) $\sum e_i^2=\sum(y_i-\hat{y}_i)^2=\sum(y_i-\hat{\beta}_0-\hat{\beta}_1x_i)^2$

$$=(10-6)^2+(10-13)^2+(20-20)^2+(20-27)^2+(40-34)^2$$

$$=110$$

$$\hat{\sigma}^2=\frac{1}{n-2}\sum e_i^2$$

$$=\frac{110}{3}$$

$$\Rightarrow \hat{\sigma}=6.055$$

(5) $t_{0.025}(3)=3.182$

① $\hat{\beta}_1\sim N(\beta_1,\frac{\sigma^2}{L_{xx}})$

$$t_1=\frac{\hat{\beta}_1-\beta_1}{\sqrt{\hat{\sigma}^2/L_{xx}}}=\frac{(\hat{\beta}_1-\beta_1)\sqrt{L_{xx}}}{\hat{\sigma}}\sim t(3)$$

$$\Rightarrow P(|t|<t_{\frac{\alpha}{2}}(3))=1-\alpha$$

$$\Rightarrow P(\hat{\beta}_1-t_{\frac{\alpha}{2}}\frac{\hat{\sigma}}{\sqrt{L_{xx}}}<\beta_1<\hat{\beta}_1+t_{\frac{\alpha}{2}}\frac{\hat{\sigma}}{\sqrt{L_{xx}}})=1-\alpha$$

$\Rightarrow$ β_1的置信度95%的置信区间为：

$$(\hat{\beta}_1-t_{0.025}\frac{\hat{\sigma}}{\sqrt{L_{xx}}},\ \hat{\beta}_1+t_{0.025}\frac{\hat{\sigma}}{\sqrt{L_{xx}}})$$

即：$(0.91, 13.09)$

② $\hat{\beta}_0\sim N(\beta_0,(\frac{1}{n}+\frac{(\bar{x})^2}{L_{xx}})\sigma^2)$

即：$\hat{\beta}_0\sim N(\beta_0, 1.1\sigma^2)$

同理可求出：

β_0的置信度95%的置信区间为：

$$(\hat{\beta}_0-t_{0.025}\sqrt{1.1}\,\hat{\sigma},\ \hat{\beta}_0+t_{0.025}\sqrt{1.1}\,\hat{\sigma})$$

即：$(-21.21, 19.21)$

(6) $r^2=\frac{SSR}{SST}=\frac{ESS}{TSS}=\frac{L_{xy}^2}{L_{xx}L_{yy}}=\frac{70^2}{10\times 600}=0.817$

(7) $SSR=ESS=\sum(\hat{y}_i-\bar{y})^2=(6-20)^2+(13-20)^2+0^2+(27-20)^2+(34-20)^2$

$$=490$$

$SSE=RSS=\sum(y_i-\hat{y}_i)^2=110$

$SST=TSS=\sum(y_i-\bar{y})^2=L_{yy}=600$

$SSR/1=490$

$SSE/(n-2)=36.67$

$F=\frac{SSR/1}{SSE/(n-2)}=13.36$

$P(F>F_\alpha(1,3))=\alpha \xRightarrow{\text{查表}} 0.025<\alpha<0.05$

故方差分析如表：

	df	SS	MS	F	p-Value
回归	1	490	490	13.36	0.025~0.05
残差	3	110	36.67		
总和	4	600			

(8) $H_0:\ \beta_1=0$，$H_1:\ \beta_1\neq 0$

$$\Rightarrow \hat{\beta}_1\sim N(0,\frac{\sigma^2}{L_{xx}})$$

令 $t=\frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2/L_{xx}}}=\frac{\hat{\beta}_1\sqrt{L_{xx}}}{\hat{\sigma}}\sim t(3)$

$$P(|t|<t_{0.025}(3))$$

$$=P(-t_{0.025}(3)\frac{\hat{\sigma}}{\sqrt{L_{xx}}}<\hat{\beta}_1<t_{0.025}(3)\frac{\hat{\sigma}}{\sqrt{L_{xx}}})$$

$$=P(-6.09<\hat{\beta}_1<6.09)$$

$$=1-\alpha=0.95$$

由 $\hat{\beta}_1=7$

故拒绝H_0，不能证明β_1在95%的置信水平上为0

$\Rightarrow$ β_1显著不为0

(9) $r=\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}=0.904$

$$t=\frac{\sqrt{n-2}\,r}{\sqrt{1-r^2}}=3.656$$

$$t_{0.025}(3)=3.182$$

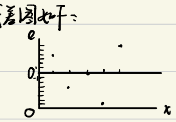
$$\Rightarrow |t|>t_{0.025}(3)$$

$\Rightarrow$ 可以认为y与x的简单回归系数显著不为0

(10) 由题：$e_1=10-6=4$，$e_2=10-13=-3$，$e_3=20-20=0$，

$e_4=20-27=-7$，$e_5=40-34=6$

残差图如下：



残差在$e=0$附近随机变化且无明显周期性与异常

但由于样本量过小，也无法得出准确判断

(11) $\hat{y}_0=-1+7\times 4.2=28.4$

$\Rightarrow$ 销售收入将达到28.4万元

由 $\hat{y}_0\sim N(\beta_0+\beta_1x_0,(\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}})\sigma^2)$ 可知：

$$y_0-\hat{y}_0\sim N(0,(1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}})\sigma^2)$$

$\Rightarrow$ 令 $t=\frac{y_0-\hat{y}_0}{\sqrt{1+\frac{1}{n}+(x_0-\bar{x})^2/L_{xx}}}\sim t(3)$

故 $P(|t|<t_{0.025}(3))=0.95$

$\Rightarrow$ y_0的置信概率95%的置信区间为

$$(\hat{y}_0-t_{0.025}(3)\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}}},\ \hat{y}_0+t_{0.025}(3)\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}}})$$

即：$(24.71, 32.01)$

Ch 3

3.3

证明：$\hat{\sigma}^2=\frac{1}{n-p-1}SSE=\frac{1}{n-p-1}e'e=\frac{1}{n-p-1}\sum e_i^2$

$$\hat{Y}=X\hat{\beta}=X(X'X)^{-1}X'Y=HY$$

$$e=Y-\hat{Y}=(I-H)Y$$

$$D(e)=Cov(e,e)=Cov((I-H)Y,(I-H)Y)=\sigma^2(I-H)$$

$$\Rightarrow E(\sum e_i^2)=\sum D(e_i)=\sum(1-h_{ii})\sigma^2$$

此时，$tr(H) = tr(X(X'X)^{-1}X')$

$= tr((X'X)^{-1}X'X)$

（由X为$n\cdot(p+1)$阶）$= tr(I_{p+1})$

$= p+1$

$\Rightarrow E(\sum e_i^2) = \sum(1-h_{ii})\sigma^2$

$= n\sigma^2 - (p+1)\sigma^2$

$\Rightarrow E(\hat{\sigma}^2) = \frac{1}{n-p-1} \cdot (n-(p+1))\sigma^2$

$= \sigma^2$

无偏得证.