例 在钢线碳含量对于电阻的效应的研究中，得到如下表所示一批数据：

含碳量 x%	0.1	0.3	0.40	0.55	0.70	0.80	0.95
电阻 $y(\Omega)$	15	18	19	21	22.6	23.8	26

(1)假设随机误差 ε 服从 $N(0,\sigma^2)$ 分布，试求 y 关于 x 的经验回归方程; (2)设 $\alpha=0.05$,试问 y 关于 x 的线性回归方程效果是否显著？ (3)在 $x=0.6$ 时，求出 y 的预测值及置信度为95%的预测区间. 三、**解：** (1)由数据表得

序号	x_i	y_i	x_i^2	y_i^2	x_iy_i
1	0.1	15	0.01	225	1.5
2	0.3	18	0.09	324	5.4
3	0.4	19	0.16	361	7.6
4	055	21	0.3025	441	11.55
5	0.70	22.6	0.49	510.76	15.82
6	0.80	23.8	0.64	566.44	19.04
7	0.95	26	0.9025	676	24.7
$\sum$	3.8	145.4	2.595	3104.2	85.61

因 $\overline{x}=\frac{1}{7}\sum_{i=1}^{7}x_i=\frac{3.8}{7}=0.543,\overline{y}=\frac{1}{7}\sum_{i=1}^{7}y_i=\frac{145.4}{7}=20.77,$

$$S_{xx}=\sum_{i=1}^{7}(x_i-\overline{x})^2=\sum_{i=1}^{7}x_i^2-7(\overline{x})^2=2.595-7\times0.543^2=0.5321$$

$$S_{xy}=\sum_{i=1}^{7}x_iy_i-7\overline{x}\times\overline{y}=85.61-7\times0.543\times20.77=6.6786$$

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=\frac{\sum_{i=1}^{7}x_iy_i-7\overline{x}\times\overline{y}}{\sum_{i=1}^{7}x_i^2-7(\overline{x})^2}=\frac{85.61-7\times0.543\times20.77}{2.595-7\times0.543^2}=12.55.$$

$\hat{a}=\overline{y}-\hat{b}\overline{x}=20.77-12.55\times 0.543=13.96$. 故所求线性回归方程为 $\hat{y}=13.96+12.55x$.

(2)假设 $H_0: b=0; H_1: b\neq 0$. 当 H_0 成立时， $T=\dfrac{\hat{b}\sqrt{S_{xx}}}{\sqrt{Q_E/n-2}}=\dfrac{S_{xy}/\sqrt{S_{xx}}}{\sqrt{Q_E/n-2}}\sim t(5)$,

因 $S_{yy}=\sum_{i=1}^{7}(y_i-\overline{y})^2=\sum_{i=1}^{7}y_i^2-7(\overline{y})^2=3104.04-7\times 20.77^2=83.8743$ ，故

$Q_E==S_{yy}-\hat{b}S_{xy}=83.8743-12.5514\times 6.6786=0.0473,$

$\hat{\sigma}^2=\dfrac{Q_E}{n-2}=\dfrac{0.0473}{5}=0.0095,$

从而 $T=\dfrac{S_{xy}/\sqrt{S_{xx}}}{\sqrt{Q_E/n-2}}=\dfrac{S_{xy}/\sqrt{S_{xx}}}{\sqrt{\hat{\sigma}^2}}=\dfrac{6.6786}{\sqrt{0.0095\times 0.5321}}=92.76>t_{0.025}(5)=2.5706,$

所以 y 关于 x 的线性回归方程效果非常显著.

(3)当 $x=0.6$ 时， y 的预测值为 $\hat{y}=13.96+12.55\times 0.6=21.4886,$

因 $T=\dfrac{y-\hat{y}}{\sqrt{Q_E/n-2}\sqrt{(1+\frac{1}{n}+\frac{(x-\overline{x})^2}{S_{xx}})}}=\dfrac{y-\hat{y}}{\sqrt{\hat{\sigma}^2(1+\frac{1}{7}+\frac{(x-\overline{x})^2}{S_{xx}})}}\sim t(5)$

当给定置信度 $1-\alpha=0.95$ 时，令 $P\{|T|\leq t_{0.025}(5)\}=1-\alpha=0.95$

因 $t_{\frac{\alpha}{2}}(n-2)\sqrt{\hat{\sigma}^2(1+\frac{1}{n}+\frac{(x-\overline{x})^2}{S_{xx}})}=t_{0.025}(5)\sqrt{0.0095(1+\frac{1}{7}+\frac{(0.6-0.543)^2}{0.5321})}=0.2686$ ，

从而 $x=0.6$ 时, 置信度为 95%的 y 的预测区间为:

$[\ \hat{y}-t_{0.025}(5)\sqrt{\hat{\sigma}^2(1+\frac{1}{7}+\frac{(x-\overline{x})^2}{S_{xx}})}\ ,\ \hat{y}+t_{0.025}(5)\sqrt{\hat{\sigma}^2(1+\frac{1}{7}+\frac{(x-\overline{x})^2}{S_{xx}})}\ ]$

$=[21.218, 21.756].$