1.小麦基本苗数 x 及有效穗数 Y（单位：万）的 5 组观察数据如下：

基本苗数 x_i	15.0	25.8	30.0	36.6	44.4
有效穗数 y_i	39.4	41.9	41.0	43.1	49.2

试求线性回归方程并用三种方法做显著性检验；若 x_0=26，求：Y_0 的 0.95 预测区间.

解:①作散点图(略);

②建模：$y_i=\beta_0+\beta_1 x_i+\varepsilon_i\sim N(\beta_0+\beta_1 x_i,\sigma^2).i=1,\cdots,n.$

③ $\bar{x}$ =30.36， $\bar{y}$ =42.92 ； $l_{xx}=\sum_{i=1}^{n}x_i^2-n\bar{x}^2=\sum_{i=1}^{n}(x_i-\bar{x})^2$ =492.912,

$l_{xy}=\sum_{i=1}^{n}y_i x_i-n\overline{xy}$ =148.704, $l_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2$ =56.588;

$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{l_{xy}}{l_{xx}}=0.302,\ \hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=33.76.$$

得经验回归方程：$\hat{y}=33.76+0.302x$.

④对 $H_0:\beta_1=0\leftrightarrow H_1:\beta_1\neq 0$ 的检验,α =0.05.

方 法 1： F 检验法(或方差分析法)

$$SSR=\hat{\beta}_1^2 l_{xx}=44.956;\ SSE=l_{yy}-\hat{\beta}_1^2 l_{xx}=11.6325;\quad MSE=SSE/(5-2)=3.87;$$

确定拒绝域: $F_{0.95}(5-2)$=10.1，W_1={ $F>F_{0.95}(5-2)$=10.1}，而 $F=SSR/MSE=11.6$.

从而 W_1 发生,故拒绝 $H_0:\beta_1=0$，接受 $H_1:\beta_1\neq 0$，即认为有线性相关性，或经验回归方程：$\hat{y}=33.76+0.302x$ 合理.

方法 2：相关系数法 $|r|=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}=\frac{|l_{xy}|}{\sqrt{l_{xx}l_{yy}}}$ =0.8904,

确定拒绝域 W_1={ $|r|>r_{0.95}(5-2)$=0.8783},从而 W_1 发生，故拒绝 $H_0:\beta_1=0$，接受 $H_1:\beta_1\neq 0$，即认为有线性相关性，或经验回归方程：$\hat{y}=33.76+0.302x$ 合理.

方法 3： t 检验法.

$$t=\hat{\beta}_1/\sqrt{MSE/l_{xx}}=0.303/\sqrt{3.877/492.912}=3.405$$

确定拒绝域 W_1={ $|t|>t_{0.975}(5-2)$=3.182},从而 W_1 发生,故拒绝 $H_0:\beta_1=0$，

接受 $H_1:\beta_1\neq 0$，即认为有线性相关性,或经验回归方程: $\hat{y}=33.76+0.302x$ 合理.

⑤⑴当 x_0=26 时，Y_0 的点估计为 $\hat{y}_0$ 点估计 $\hat{y}_0=33.76+0.302\times 26$ =41.6047，

⑵$\Delta=t_{0.975}\sqrt{[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}]MSE}=3.182\sqrt{[1+\frac{1}{5}+\frac{(26-30.36)^2}{492.912}]\times 3.877}$ =6.973,

从而预测区间为：(41.6047－6.973，41.6047＋6.973)=(34.63,48.58).

2．北碚大红番茄果实横径 x 单位：ｃｍ）与果重 Y (单位：ｇ）的观察数据如下：

果实横径 x_i	10	9.6	9.2	8.9	8.5	8.0	7.8	7.7	7.4	7.0
果重 Y_i	140	132	130	121	116	108	105	106	95	90

试求线性回归方程并用三种方法做显著性检验.

分析:①作散点图(略);②建模: $y_i=\beta_0+\beta_1x_i+\varepsilon_i\sim N(\beta_0+\beta_1x_i,\sigma^2).i=1,\cdots,n.$

③ $\bar{x}$＝8.41，$\bar{y}$＝114.3 ； $l_{xx}=\sum_{i=1}^{n}x_i^2-n\bar{x}^2=\sum_{i=1}^{n}(x_i-\bar{x})^2$ =8.869,

$l_{xy}=\sum_{i=1}^{n}y_ix_i-n\bar{x}\bar{y}$=145.67, $l_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2$ =2426.1;

$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{l_{xy}}{l_{xx}}$=16.425, $\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}$=-23.83

得经验回归方程: $\hat{y}=-23.83+16.425x$.

④对 $H_0:\beta_1=0\leftrightarrow H_1:\beta_1\neq 0$ 的检验,α =0.05.

方法 1：F 检验法(或方差分析法)

$SSR=\hat{\beta}_1^2l_{xx}$=2392.575; $SSE=l_{yy}-\hat{\beta}_1^2l_{xx}$=33.525; $MSE=SSE/(10-2)$=4.2;

确定拒绝域: $F_{0.95}(10-2)$=19.4, W_1={$F>F_{0.95}(5-2)$=19.4}, 而 $F=SSR/MSE$=569.66, 从而 W_1 发生，故拒绝 $H_0:\beta_1=0$，接受 $H_1:\beta_1\neq 0$，即认为有线性相关性，或经验回归方程: $\hat{y}=-23.83+16.425x$ 合理.

方法 2：相关系数检验法

$$|r|=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}=\frac{|l_{xy}|}{\sqrt{l_{xx}l_{yy}}}=0.993,$$

确定拒绝域 W_1={$|r|>r_{0.95}(10-2)$=0.6319},从而 W_1 发生,故拒绝 $H_0:\beta_1=0$，

接受$H_1: \beta_1 \neq 0$,即认为有线性相关性,或经验回归方程:$\hat{y}=-23.83+16.425x$合理.

方法 3： t检验法

$$t=\hat{\beta}_1\Big/\sqrt{MSE/l_{xx}}=16.425\Big/\sqrt{4.2/8.869}=23.868$$

确定拒绝域$W_1=\{|t|>t_{0.975}(10-2)=2.306\}$,从而$W_1$发生,故拒绝$H_0: \beta_1=0$，

接受$H_1: \beta_1 \neq 0$即认为有线性相关性，或经验回归方程：$\hat{y}=-23.83+16.425x$合理.

3.害虫的发生与气象条件有一定的关系.某地观测1964年－1973年7月下旬的温雨系数x(雨量/温度)和大豆第二代 造桥虫发生量Y(每百株虫数)的数据如下:

年份	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973
x_i	1.58	9.98	9.42	1.25	0.30	2.41	11.01	1.85	6.04	5.92
Y_i	180	28	25	117	165	175	40	160	120	80

试求线性回归方程并用三种方法做显著性检验.

分析:①作散点图(略); ②建模: $y_i=\beta_0+\beta_1 x_i+\varepsilon_i \sim N(\beta_0+\beta_1 x_i,\sigma^2). i=1,\cdots,n.$

③ $\bar{x}$＝4.979，$\bar{y}$＝109 ； $l_{xx}=\sum_{i=1}^{n}x_i^2-n\bar{x}^2=\sum_{i=1}^{n}(x_i-\bar{x})^2$ =146.92,

$l_{xy}=\sum_{i=1}^{n}y_i x_i-n\overline{xy}$ =-2073.07, $l_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2$ =234538;

$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{l_{xy}}{l_{xx}}=-14.11,\quad \hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=179.256,$$

得经验回归方程: $\hat{y}=179.256-14.11x$.

④对$H_0: \beta_1=0 \leftrightarrow H_1: \beta_1 \neq 0$的检验,$\alpha$=0.05

方法 1： F检验法(或方差分析法)

$SSR=\hat{\beta}_1^2 l_{xx}$=29252.98; $SSE=l_{yy}-\hat{\beta}_1^2 l_{xx}$=5285.02;

$MSE=SSE/(10-2)$=660.6275; 确定拒绝域: $F_{0.95}(10-2)$=19.4,

$W_1=\{F>F_{0.95}(5-2)=19.4\}$, 而$F=SSR/MSE$=44.28, 从而$W_1$发生，故拒绝

$H_0: \beta_1=0$，接受$H_1: \beta_1 \neq 0$，即认为有线性相关性，或认为经验回归方程:

$\hat{y}=179.256-14.11x$合理.

方法 2：相关系数检验法

$$|r|=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}=\frac{|l_{xy}|}{\sqrt{l_{xx}l_{yy}}}=0.92,$$

确定拒绝域$W_1=\{|r|>r_{0.95}(10-2)=0.6319\}$,从而$W_1$发生,故拒绝$H_0:\beta_1=0$，接受

$H_1:\beta_1\neq 0$，即认为有线性相关性,或认为经验回归方程：$\hat{y}=179.256-14.11x$合理.

方法 3： t检验法

$$t=\frac{\hat{\beta}_1}{\sqrt{MSE/l_{xx}}}=\frac{14.11}{2.12}=6.656;$$

确定拒绝域$W_1=\{|t|>t_{0.975}(10-2)=2.306\}$,从而$W_1$发生,故拒绝$H_0:\beta_1=0$，

接受$H_1:\beta_1\neq 0$，即认为有线性相关性，或经验回归方程：$\hat{y}=179.256-14.11x$合理.

4.某地国民生产总值 Y(单位:亿元)与基本建设投资 x(单位:亿元)的年度统计数字如下:

x_i	191.72	203.66	223.11	242.82	265.45	297.62	322.00	352.41
y_i	15.53	12.92	17.62	14.21	16.90	25.58	28.00	32.47

试求线性回归方程并用三种方法做显著性检验.

解　同上面习题解法完全一样,略.