10.1.8 Since $x = 23$ and $n - x = 301$ are lager than 5, we can use the normal approximation.

$$z = \frac{23 + 0.5 - 324 * 0.1}{\sqrt{324 * 0.1 * 0.9}} = -1.648 \quad \textbf{(+2 points)}$$
$$\text{P-value} = \Phi(-1.648) = 0.0497 \quad \textbf{(+3 points)}$$

As $z_{0.01} = 2.326$, 99% upper confidence interval is

$$(0, \frac{23}{324} + \frac{2.326}{324}\sqrt{\frac{23(324-23)}{324}}) = (0, 0.1042) \quad \textbf{(+3 points)}$$

Since P-value$= 0.0497 > 0.01$, the screening test is not acceptable. **(+2 points)**

10.2.2 $x = 261$, $n - x = 41$, $y = 401$ and $n - y = 53$ are all larger than 5.

$$\hat{p}_A - \hat{p}_B = \frac{261}{302} - \frac{401}{454} = -0.019$$
$$\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{302} + \frac{\hat{p}_B(1-\hat{p}_B)}{454}} = 0.0248$$

(a) As $z_{0.005} = 2.576$, a two-sided 99% confidence interval is

$$(-0.019 - 2.576 * 0.0248, -0.019 + 2.576 * 0.0248) = (-0.083, 0.045) \quad \textbf{(+2 points)}$$

(b) As $z_{0.05} = 1.645$, a two-sided 90% confidence interval is

$$(-0.019 - 1.645 * 0.0248, -0.019 + 1.645 * 0.0248) = (-0.060, 0.022) \quad \textbf{(+2 points)}$$

(c) As $z_{0.05} = 1.645$, an one-sided 95% confidence interval is

$$(-1, -0.019 + 1.645 * 0.0248) = (-1, 0.022) \quad \textbf{(+2 points)}$$

(d) For testing $H_0 : p_A = p_B$ versus $H_A : p_A \neq p_B$

$$\hat{p} = \frac{261 + 401}{302 + 454} = 0.8757$$
$$z = \frac{-0.019}{\sqrt{0.8757(1 - 0.8757)(1/302 + 1/454)}} = -0.776 \quad \textbf{(+2 points)}$$
$$\text{P-value} = 2\Phi(-0.776) = 0.4378 \quad \textbf{(+2 points)}$$

10.2.12 Let probabilities of the original p_A and the modification p_B.
Test $H_0 : p_A \geq p_B$ versus $H_A : p_A < p_B$. **(+3 points)**

$$\hat{p}_A = \frac{22}{542},\ \hat{p}_B = \frac{64}{601},\ \hat{p} = \frac{22+64}{542+601} = \frac{86}{1143}$$

$$z = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)}} = -4.217 \quad \textbf{(+3 points)}$$

$$\text{P-value} = \Phi(-4.217) \approx 0$$

Since P-value is smaller than $\alpha = 0.05$, the null hypothesis is rejected then there is evidence that the modifications attracted more customers. **(+4 points)**

10.3.6 Degree of freedom=2. $\chi^2_{0.5} = 5.99$. **(+3 points)** Expected cell frequency is 200. **(+2 points)** $\chi^2 = \frac{(200-225)^2}{200} + \frac{(200-223)^2}{200} + \frac{(200-152)^2}{200} = 17.29$. Therefore, we reject null hypothesis. **(+5 points)**

10.3.10 Degree of freedom=2. $\chi^2_{0.5} = 5.99$. **(+3 points)** Expected cell frequency is $205 \times \frac{1}{3} = \frac{205}{3}$. **(+2 points)** $\chi^2 = \frac{(83-68.3)^2}{68.3} + \frac{(75-68.3)^2}{68.3} + \frac{(47-68.3)^2}{68.3} = 10.46$. Therefore, we reject null hypothesis. **(+5 points)**

10.3.12 The cumulative probability of failure to 24 hours $= 0.705$, 48 hours $= 0.812$, 72 hours $= 0.865$, $\lim_{x\to\infty}[x \text{ hours}] = 1$. **(+2 points)** Therefore, $e_1 = 88.15$, $e_2 = 13.28$, $e_3 = 6.69$, $e_4 = 16.88$. **(+2 points)** Degree of freedom=3. **(+1 points)** $\chi^2 = \frac{(88.15-12)^2}{88.15} + \frac{(13.28-53)^2}{13.28} + \frac{(6.68-39)^2}{6.68} + \frac{(16.88-21)^2}{16.88} = 341.65$. **(+5 points)** Therefore, we reject null hypothesis.

10.4.2 H_0 : two factors are independent, H_A : not H_0. We have the following data.

	No Fertilizer	Fertilizer 1	Fertilizer 2	
Dead	48	71	63	$x_{1\cdot} = 182$
Slow Growth	111	89	95	$x_{2\cdot} = 295$
Mediam Growth	186	174	181	$x_{3\cdot} = 541$
Strong Growth	142	181	190	$x_{4\cdot} = 513$
	$x_{\cdot 1} = 487$	$x_{\cdot 2} = 515$	$x_{\cdot 3} = 529$	$n = 1531$

Therefore, the expected cell frequencies are

	No Fertilizer	Fertilizer 1	Fertilizer 2
Dead	57.893	61.221	62.886
Slow Growth	93.837	99.233	101.930
Mediam Growth	172.088	181.982	186.929
Strong Growth	163.182	172.564	177.255

Degree of freedom is $(4-1)(3-1) = 6$. The Pearson chi-square statistic is $X^2 = 13.659$ so the p-value is $\mathbf{Pr}(X_6^2 \geq 13.659) = 0.0337$. Therefore, we can conclude that the seedlings growth pattern is different for the different growing conditions.

10.4.6 We know

$$X^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(x_{ij} - e_{ij})^2}{e_{ij}}$$

with

$$e_{ij} = \frac{x_{i\cdot}x_{\cdot j}}{n} = \frac{(x_{i1} + x_{i2})(x_{1j} + x_{2j})}{n} \quad \text{and} \quad n = x_{11} + x_{12} + x_{21} + x_{22}.$$

For each i, j, we have

$$\frac{(x_{ij} - e_{ij})^2}{e_{ij}} = \frac{n^2}{n^2} \times \frac{(x_{ij} - \frac{x_{i\cdot}x_{\cdot j}}{n})^2}{\frac{x_{i\cdot}x_{\cdot j}}{n}} = \frac{(nx_{ij} - x_{i\cdot}x_{\cdot j})^2}{nx_{i\cdot}x_{\cdot j}}$$

Using $n = x_{11} + x_{12} + x_{21} + x_{22}$, we have

$$\frac{(nx_{ij} - x_{i\cdot}x_{\cdot j})^2}{nx_{i\cdot}x_{\cdot j}} = \frac{((x_{11} + x_{12} + x_{21} + x_{22})x_{ij} - x_{i\cdot}x_{\cdot j})^2}{nx_{i\cdot}x_{\cdot j}} \tag{1}$$

For any $i, j \in \{1, 2\}$, we can check that

$$(x_{11} + x_{12} + x_{21} + x_{22})x_{ij} - x_{i\cdot}x_{\cdot j} = x_{11}x_{22} - x_{12}x_{21}.$$

That is, RHS of (1) becomes

$$\frac{(x_{11}x_{22} - x_{12}x_{21})^2}{nx_{i\cdot}x_{\cdot j}}$$

Since we have

$$\frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{2} \frac{1}{x_{i\cdot}x_{\cdot j}} = \frac{x_{1\cdot}x_{\cdot 1} + x_{1\cdot}x_{\cdot 2} + x_{2\cdot}x_{\cdot 1} + x_{2\cdot}x_{\cdot 2}}{nx_{1\cdot}x_{\cdot 1}x_{2\cdot}x_{\cdot 2}} = \frac{n}{x_{1\cdot}x_{\cdot 1}x_{2\cdot}x_{\cdot 2}},$$

we conclude

$$X^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(x_{ij} - e_{ij})^2}{e_{ij}} = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(x_{11}x_{22} - x_{12}x_{21})^2}{nx_{i\cdot}x_{\cdot j}} = \frac{n(x_{11}x_{22} - x_{12}x_{21})^2}{x_{1\cdot}x_{\cdot 1}x_{2\cdot}x_{\cdot 2}}.$$

10.4.10 We have below expected cell frequencies.

Type	Severe	Medium	Minor	
A	$x_{11} = 9$ $e_{11} = 8.14$	$x_{12} = 17$ $e_{12} = 13.09$	$x_{13} = 31$ $e_{13} = 35.77$	$x_{1\cdot} = 57$
B	$x_{21} = 4$ $e_{21} = 7.00$	$x_{22} = 9$ $e_{22} = 11.25$	$x_{23} = 36$ $e_{23} = 30.75$	$x_{2\cdot} = 49$
C	$x_{31} = 15$ $e_{31} = 12.86$	$x_{32} = 19$ $e_{32} = 20.66$	$x_{33} = 56$ $e_{33} = 56.48$	$x_{3\cdot} = 90$
	$x_{\cdot 1} = 28$	$x_{\cdot 2} = 45$	$x_{\cdot 3} = 123$	$n = x_{\cdot\cdot} = 196$

The Pearson chi-square statistic is

$$X^2 = \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{(x_{ij} - e_{ij})^2}{e_{ij}} = 5.024.$$

The degree of freedom is $(3-1) \times (3-1) = 4$. Thus, the $p-$value is $P(\chi_4^2 \geq 5.024) = 0.285$. Since $0.285 > 0.05$, the null hypothesis of independence is plausible. (There is no sufficient evidence to conclude that the three types of asphalt are different with respect to cracking)

10.7.20 Let the chances of success for patients with condition A and B be denoted by P_A and P_B, respectively. Then $\hat{P_A} = 56/94$ and $\hat{P_B} = 64/153$. For solving (c), define x_{ij} and e_{ij} as :

$$\begin{array}{cccc} & \text{success} & \text{failiure} & \\ A & x_{11} = 56 & x_{12} = 38 & x_{1.} = 94 \\ B & x_{21} = 64 & x_{22} = 89 & x_{2.} = 153 \\ & x_{.1} = 120 & x_{.2} = 127 & x_{..} = 247 \end{array}$$

$$\begin{array}{cccc} & success & failiure & \\ A & e_{11} = 45.67 & e_{12} = 48.33 & 94 \\ B & e_{21} = 74.33 & e_{22} = 78.64 & 153 \\ prob & x_{.1}/x_{..} = 0.486 & x_{.2}/x_{..} = 0.514 & \end{array}$$

(a) The hypothesis is $H_0 : P_A \leq 0.5$ versus $H_A : P_A > 0.5$. **(+5 points)**

$$z = \frac{x - 0.5 - np_0}{\sqrt{np_0(1-p_0)}} = 1.753,$$

where $n = 94$ and $p_0 = 0.5$. And the p-value is $1 - \phi(1.753) = 0.04 < 0.05$. So, H_0 is rejected at the siginificant level $\alpha = 0.05$. There is sufficient evidence to conclude that the chance of success for patients with condition A is better than 50%. **(+5 points)**

(b) Set $\alpha = 0.01$. Then the 99% confidence interval is $(0.012, 0.344)$ which is from

$$\hat{P_A} - \hat{P_B} \pm z_{\alpha/2}\sqrt{\frac{\hat{P_A}(1-\hat{P_A})}{n_A} + \frac{\hat{P_B}(1-\hat{P_B})}{n_B}}.\textbf{(+10 points)}$$

(c) Set $\alpha = 0.05$. The hypothesis is H_0 : the success probabilities are same for patients with condition A and with condition B versus H_A : not H_0. **(+5 points)** Use 10.4.6, then

$$X^2 = \sum_{1\leq i,j\leq 2} \frac{(x_{ij} - e_{ij})^2}{e_{ij}} = \frac{x_{..}(x_{11}x_{22} - x_{12}x_{21})^2}{x_{1.}x_{.1}x_{2.}x_{.2}} = 7.339, \quad \text{df} = 1.$$

and the p-value is $P(\chi_1^2 \geq 7.339) = 0.007 < 0.05$. Therefore, the null hypothesis is rejected at the significant level and the success probabilities are different . **(+5 points)**